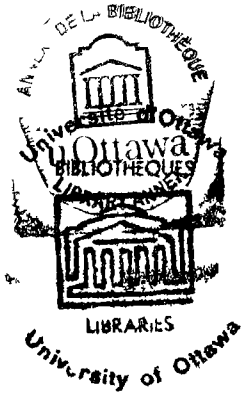
THE INFLUENCE OF SOCIAL DESIRABILITY RESPONSE SET AND OF CONFORMITY ON THE CORRESPONDENCE BETWEEN RESPONSES TO SELF-REPORT AND ERROR-CHOICE PERSONALITY INVENTORIES

by Paul E. Sayer

Thesis presented to the School of Graduate Studies of the University of Ottawa as partial fulfillment of the requirements for the degree of Master of Arts



Ottawa, Canada, 1972

ACKNOWLEDGMENTS

This thesis was prepared under the supervision of Henry P. Edwards, Ph.D., ~~Assistant~~ Associate Professor at the Faculty of Psychology. The writer is appreciative of the guidance received from the members of the advisory team, William F. Barry, Ph.D., and Georges Sarrazin, Ph.D.

Gratitude is expressed to Barry Bragg, Ph.D., University of Toronto, for his generous sharing of laboratory facilities, and to Joan Sayer for her assistance in the administration and scoring of tests.

The writer extends thanks to the volunteer subjects.

CURRICULUM STUDIORUM

Paul E. Sayer was born February 6, 1943, in Ottawa, Ontario. He received the Bachelor of Arts degree from St. Patrick's College, Ottawa, in 1964.

TABLE OF CONTENTS

Chapter page

INTRODUCTION. ix

I.- REVIEW OF THE LITERATURE. 1
1. The Social Desirability Response Set Controversy 2
2. The Error-Choice Technique 13
3. SD Response Set and Error-Choice Inventory 20
4. Conformity and Error-Choice Inventory 26
5. Summary and Experimental Hypotheses 31

II.- EXPERIMENTAL DESIGN 34
1. The Sample 35
2. The General Method of the Experiment 35
3. The Tools and Related Procedures 38
A. The Measure of SD Response Set 38
B. The Measure of Conformity 42
C. The Self-Report Inventory 51
D. The Error-Choice Inventory 51
4. Statistical Analysis 54
5. The Specific Hypotheses 55

III.- PRESENTATION AND DISCUSSION OF RESULTS 58
1. Results of the Tests of Significance 58
2. Discussion of Results 64

SUMMARY AND CONCLUSIONS 78

BIBLIOGRAPHY. 80

Appendix

1.- THE INVENTORY MEASURING SD RESPONSE SET 84

2.- THE KEYING AND THE CORRESPONDING MMPI BOOKLET NUMBER OF THE ITEMS OF THE "INTRODUCTORY INVENTORY". 89

3.- INSTRUCTIONS TO SUBJECTS IN THE PROCEDURE FOR MEASURING CONFORMITY. 91

4.- THE CRITICAL ITEMS IN THE PROCEDURE FOR MEASURING CONFORMITY. 96

Appendix page

5.- INSTRUCTIONS PLUS ANSWER SHEET FOR THE EC INVENTORY. 111

6.- THE POST-EXPERIMENT QUESTIONNAIRE 113

7.- MEAN SR-EC CORRESPONDENCE SCORES FOR THE FACTOR COMBINATIONS. 115

8.- RAW DATA. 117

9.- ABSTRACT OF The Influence of Social Desirability Response Set and of Conformity on the Correspondence Between Responses to Self-Report and Error-Choice Personality Inventories 123

LIST OF TABLES

Table page

I.- The Changes in Item Endorsement Frequencies to Form the Answer Choices for the EC Inventory. 52

II.- Distribution and Correlation of SD Response Set Scores and Conformity Scores 59

III.- Three-Way, Fixed Effects Analysis of Variance of SR-EC Correspondence Scores for Different Levels of SD Response Set and Conformity and Different Orders of Administration 61

IV.- Analysis of Main Effects Variance of SR-EC Correspondence Scores, Using a Pooled Error Term 62

V.- Analysis of Variance for Simple Effects of Orders of Administration for Each Level of SD Response Set 74

VI.- The Keying and the Corresponding MMPI Booklet Number of the Items of the "Introductory Inventory" 90

VII.- Mean SR-EC Correspondence Scores for the Factor Combinations of the Analysis of Variance Design 116

VIII.- Subjects' Scores for SD Response Set, Conformity, and SR-EC Correspondence 118

LIST OF FIGURES

Figure page

1.- Diagram of the Laboratory for Measuring Conformity-Proneness. 43

2.- Diagram of the Answering Console Used in the Procedure for Measuring Conformity-Proneness. . . 45

3.- Diagram of the Analysis of Variance Design. 56

4.- Diagram of the First Critical Item in the Procedure for Measuring Conformity-Proneness. . . 97

5.- Diagram of the Second Critical Item in the Procedure for Measuring Conformity-Proneness. . . 98

6.- Diagram of the Third Critical Item in the Procedure for Measuring Conformity-Proneness. . . 99

7.- Diagram of the Fourth Critical Item in the Procedure for Measuring Conformity-Proneness. . . 100

8.- Diagram of the Fifth Critical Item in the Procedure for Measuring Conformity-Proneness. . . 101

9.- Diagram of the Sixth Critical Item in the Procedure for Measuring Conformity-Proneness. . . 102

10.- Diagram of the Seventh Critical Item in the Procedure for Measuring Conformity-Proneness. . . 103

11.- Diagram of the Eighth Critical Item in the Procedure for Measuring Conformity-Proneness. . . 104

12.- Diagram of the Ninth Critical Item in the Procedure for Measuring Conformity-Proneness. . . 105

13.- Diagram of the Tenth Critical Item in the Procedure for Measuring Conformity-Proneness. . . 106

14.- Diagram of the Eleventh Critical Item in the Procedure for Measuring Conformity-Proneness. . . 107

Figure page

15.- Diagram of the Twelfth Critical Item in the Procedure for Measuring Conformity-Proneness. . . 108

16.- Diagram of the Thirteenth Critical Item in the Procedure for Measuring Conformity-Proneness. . . 109

17.- Diagram of the Fourteenth Critical Item in the Procedure for Measuring Conformity-Proneness. . . 110

INTRODUCTION

It has been claimed that self-report personality inventories offer advantages of efficiency and objectivity in psychological assessment. Inventories are widely used, but their popularity exists in spite of considerable criticism with regard to their validity in measuring a testee's actual behaviors or traits. During the last fifteen years a major objection has been that validity is sabotaged by the operation of unconscious response sets, chiefly those of social desirability and acquiescence. Various modifications of the self-report inventory have been attempted but none has been notably successful in answering the criticism, so that the search for more effective means of personality assessment has continued.

A psychometric method termed the error-choice technique recently was adapted for use with personality inventories, and it was proposed that the technique might have greater validity than self-report inventory for revealing the traits of the respondent. In particular, it was argued that the error-choice format of personality inventory might avoid the inaccuracy contributed by the response set of social desirability.

The beginnings of experimental investigation were made, and theories were put forward to account for the relation between self-report and error-choice responses. Ultimately needed would be large-scale studies of validity, using rigorous

criterion measures. Before such extensive undertakings, however, it was appropriate to submit to experimental examination the theoretical aspects of the technique. The present report describes an investigation of the influence of social desirability response set on the correspondence between responses to self-report and error-choice inventories.

In addition, the study investigates the theory that a conformity motivation underlies the relation between self-report and error-choice inventories. The theory carried the implication that the error-choice technique of personality assessment might be applicable only to highly conforming individuals. Such a limitation of applicability would be an important practical consideration in evaluating the merits of the proposed method of personality assessment.

The experiment also examines the order of administration of the self-report and error-choice inventories as a possible influence on the correspondence between their responses.

The report begins with a review of the literature pertaining to social desirability response set as a factor in the validity of self-report inventories. The theory and development of the error-choice technique of psychological measurement are reported next. Then there is a discussion of theory relating the technique to the response set of social desirability and to the personality dimension of conformity.

In Chapter II the experimental design and the procedure of the project are elaborated, and the tools are described. The methods of statistical analysis are identified, and the hypotheses are then presented in the null form.

The last chapter presents and discusses the results of the statistical analysis, emphasizing their meaning in relation to the theories advanced by the proponents of the error-choice format for personality inventory. Suggestions are made with regard to further research.

CHAPTER I

REVIEW OF THE LITERATURE

There is a substantial body of literature which presents the argument that unconscious response sets or response styles reduce the validity of self-report personality inventories. There are other authors who argue that response sets or styles are of little or no consequence. It is an area of controversy, with problems of definition and measurement. Of the two response sets most prominent in the literature, social desirability and acquiescence, the former has been supported with the more precise definition and the more consistent evidence. The present chapter briefly reviews major arguments in the debate about social desirability response set.

Next the theory and development of the error-choice technique are described, including the use of the technique with personality inventories. The relation of social desirability response set to the error-choice format of personality inventory is discussed. Then a rationale hypothesized to underlie inventory responses is reported and linked to the personality dimension of conformity-proneness. These theoretical formulations are reduced to the experimental hypotheses, which conclude the chapter.

1. The Social Desirability Response Set Controversy.

The concept of set is common in psychology; it denotes a tendency to respond in a particular direction. Inasmuch as there exist individual differences, then subjects have predisposing tendencies that, in part, predetermine or bias their responses to stimuli. Understood thus, set is nearly equivalent to personality or to personality trait for some theorists,[1] and sets can be expected in the responding to personality inventories. However, a more restricted meaning is generally intended for response sets as they pertain to inventories; the meaning is illustrated in the definition of response styles below.

Response styles are considered equivalent to sets by some authors, but others have made a distinction. Jackson and Messick, two of the leading participants in the controversy, defined response styles as:

> expressive consistencies in the behavior of respondents which are relatively enduring over time, with some degree of generality beyond a particular test performance to responses both in other tests and in non-test behavior, and usually reflected in assessment situations by consistencies in response to item characteristics other than specific content.[2]

---

1 Richard K. McGee, "Response Set in Relation to Personality: An Orientation," in I.A. Berg (ed.), Response Set in Personality Assessment, Chicago, Aldine, 1967, p. 1-2.

2 Douglas N. Jackson and Samuel Messick, "Response Styles and the Assessment of Psychopathology," in Samuel Messick and John Ross (eds.), Measurement in Personality and Cognition, New York, Wiley, 1962, p. 134.

That definition makes response styles comparable to personality traits in that the expressive consistencies operate in non-test behavior as well as in test responses. Other authors seem to limit the generality of response styles to test situations.[3] However, an essential element of the concept is commonly agreed upon: response set and response style usually mean a tendency to select some response category disproportionately often, independently of the item content.[4] ("Response set" and "response style" will be used interchangeably as equivalent terms in the report.)

Those who emphasize the importance of response sets or styles question the assumptions that a subject's response to an inventory item is determined essentially by the particular content of the item, and that his response is self-descriptive in terms of the item content. If response styles do operate, at the least the response may not be 100 per cent content-determined. At the extreme, the response may be determined essentially by noncontent factors, including response styles. In either case response styles and item content are confounded sources of variance in inventory scores. If such noncontent factors operated with equal strength in all

---

3 Benjamin Kleinmuntz, Personality Measurement, Homewood, Ill., Dorsey, 1967, p. 248.

4 Leonard G. Rorer, "The Great Response-Style Myth," Psychological Bulletin, Vol. 63, No. 3, 1965, p. 134.

inventory items, they would account for much of the variance in all the different inventory scales which purport to be measuring distinct traits. The validity of the scales would be thus weakened: the scales would be weakened in their correlation with criterion variables, and they would be weakened in discriminant validity because of the considerable intercorrelation between the supposedly distinct scales.[5] Inventories would not yield valid measures for the separate personality dimensions which they claimed to be measuring.

The response set of social desirability[6] has a prominent place in the literature, and it is one of the most systematically researched and developed. Edwards,[7] its leading proponent, conceived of SD response set as the tendency to give socially desirable responses in self-description, and he considered it a general personality trait which should operate in various self-report personality scales. His reasoning took the following course. For every statement in a personality inventory, a social desirability scale value[8] can be

---

5 Douglas N. Jackson, Personality Research Form Manual, Goshen, N.Y., Research Psychologists, 1967, p. 14.

6 The abbreviation SD will be used for "social desirability" throughout the report.

7 Allen L. Edwards, "The Social Desirability Variable: A Broad Statement," in I.A. Berg (ed.), Response Set in Personality Assessment, Chicago, Aldine, 1967, p. 46-47.

8 The abbreviation SDSV will be used for "social desirability scale value" throughout the report.

obtained. The SDSV is the average rating when a large number of judges have rated a statement on a rating scale, judging how socially desirable or undesirable they would consider the content of the statement if it were used to describe another person. Such SDSVs were found to be highly reliable across comparable groups of judges, and highly reliable when judges repeated their ratings.[9]

Edwards found that for personality statements the probability of endorsement increases linearly with the independently derived SDSV; i.e., the proportion of TRUE responses for a group of subjects is correlated with the SDSVs of the statements.[10] The original correlation was .87 and subsequent large studies found correlations of .89 and .90.[11] Edwards took these findings to indicate the influence of SD response set in self-report responses to personality inventories, and, therefore, the influence of SD response set on inventory trait scores. A study by Rosen,[12] which did

---

9 Allen L. Edwards, The Measurement of Personality Traits by Scales and Inventories, New York, Holt, Rinehart and Winston, 1970, p. 91.

10 --------, "The Relationship Between the Judged Desirability of a Trait and the Probability That the Trait Will Be Endorsed," Journal of Applied Psychology, Vol. 37, No. 2, 1953, p. 90-93.

11 --------, The Measurement of Personality Traits by Scales and Inventories, p. 106-107.

12 Ephraim Rosen, "Self-Appraisal, Personal Desirability, and Perceived Social Desirability of Personality Traits," Journal of Abnormal and Social Psychology, Vol. 52, No. 2, 1956, p. 151-158.

not employ SDSVs, suggested the same conclusions. He had college sophomores answer the MMPI under three conditions: (1) as self-appraisal, (2) answering TRUE if they thought the item personally desirable, and (3) answering TRUE if they thought society would consider the item desirable. The group norms of personal and social desirability were found to be similar to the endorsement norms. For the group, the correlation between endorsement and personal rating of desirability was .87.

When the SDSVs of personality statements are known, a socially desirable response may be defined as a TRUE response to an item with an SDSV above the neutral point on the rating scale or as a FALSE response to an item with an SDSV below the neutral point. In any personality scale, the keyed response to the items is the response that indicates the presence of the trait the scale was designed to measure. If all the keyed responses of a trait scale were also socially desirable responses, then the trait responses and the SD responses would be completely confounded. Scores on the scale might be measures of individual differences in the trait, or they might reflect individual differences in SD response set, the tendency to give socially desirable responses in self-description.[13] There would be similar confounding of

---

13 Edwards, The Measurement of Personality Traits by Scales and Inventories, p. 124.

the meaning of scores if all the keyed responses of a trait scale were socially undesirable responses.

In order to investigate the common-factor variances of various personality trait scales and of the tendency to give socially desirable responses, Edwards, Diers and Walker[14] performed a principal-component factor analysis of fifty-eight MMPI trait scales, one of which was constructed to measure SD response set. The latter was termed an SD scale. The ten factors extracted were rotated orthogonally by Varimax rotation. The first factor accounted for thirty-eight per cent of the total variance and forty-three per cent of the common variance, and the SD scale had its highest loading, .97, on this factor. Nine other MMPI scales had loadings (either negative or positive) of .90 and above on this factor. Altogether, twenty-four other MMPI scales had loadings of |.70| and above on the factor. The first factor was interpreted as measuring individual differences in rate of socially desirable responding. It was concluded that several MMPI scales were measuring the same trait as the SD scale, and not the separate personality dimensions that their names implied. On the basis of such experimental findings Edwards argued that SD response style severely limits the

---

[14] Allen L. Edwards, Carol J. Diers and Jerald N. Walker, "Response Sets and Factor Loadings on Sixty-One Personality Scales," Journal of Applied Psychology, Vol. 46, No. 3, 1962, p. 220-225.

ability of personality inventories to provide measures of individual differences in various personality traits.

However, other authors have disputed the contention that SD response set invalidates self-report personality scale scores. Dahlstrom[15] questioned the meaning that Edwards ascribed to the correlation between SDSV and endorsement. If desirability and item content are confounded determinants of response, Dahlstrom argued, it cannot be concluded that a subject's responding reflects SD response set; the subject's responding may represent accurate self-description, notwithstanding its congruence with socially desirable responding. In addition he noted the "circularity" of the SDSV-endorsement correlation. When judges were asked to rate the social desirability of items they might have rated some items highly favorable because the statements were descriptive of themselves. If such statements were endorsed by subjects in self-report, it would not be because of their socially desirable value. According to Dahlstrom the factorial identification of response styles as sources of variance does not necessarily deny that valid inferences about non-test behavior of subjects can be made from test scores.

---

15 W. Grant Dahlstrom, "Commentary: The Roles of Social Desirability and Acquiescence in Responses to the MMPI," in Samuel Messick and John Ross (eds.), Measurement in Personality and Cognition, New York, Wiley, 1962, p. 157-168.

Rorer[16] also rejected the conclusion from factor analytic studies of MMPI trait scores that response styles are the source of the variance. He noted the arbitrariness of labelling the factors in terms of response styles rather than in terms of the criterion groups which the MMPI scales were designed to discriminate. Rorer emphasized that, although factor analytic findings might be consistent with a response style interpretation, they did not prove the existence of styles. There was no basis for an assumption that individuals were not answering truthfully to the content of items.

Among the studies cited by Rorer in his review was one by Block.[17] Block performed a factor analysis of MMPI responses and extracted the two factors often interpreted as the response styles of social desirability and acquiescence. He then obtained Q-sort descriptions of individuals scoring high and low on the two factors. These descriptions corresponded to what one would expect on the basis of the content of the scales marking the factor positions. Block concluded the factors were not interpretable in terms of response styles.

---

16 Rorer, "The Great Response-Style Myth," p. 129-156.

17 Jack Block, The Challenge of Response Sets, New York, Appleton-Century-Crofts, 1965, quoted by Rorer, "The Great Response-Style Myth," p. 144.

Hanley[18] acknowledged the large covariance between SDSV and item endorsement but he maintained that it does not prevent high concurrent validity for diagnostic scales. His data showed very adequate discrimination between criterion groups by means of the MMPI Pd scale: of sixty subjects, fifty-four were correctly classified.

Alker[19] argued that SD response set contributes to the effectiveness of inventories. He contended that the operation of SD set allows the inventory to discriminate between genuine socially desirable behavior (which he termed effective, coping behavior) and a defensive concern with appearing socially desirable (termed maladaptive, defensive behavior). Items of low SDSV were found to be most effective in the discrimination, items of rather neutral SDSV less so. It was thus argued that minimizing the influence of SD set by using neutral items reduces the validity of the inventory.

Attempts have been made to construct inventories which control for response styles. Each set of alternatives in some forced-choice inventories has been balanced for social desirability. But after reviewing relevant studies,

---

18 Charles Hanley, "Individual Responses and Social Desirability," Journal of Applied Psychology, Vol. 51, No. 2, 1967, p. 167-169.

19 Henry A. Alker, "Coping, Defense and Socially Desirable Responses," Psychological Reports, Vol. 22, No. 3, 1968, p. 985-988.

Braun[20] concluded that SD response set is not controlled in the forced-choice format. The weakness of the forced-choice technique as a means of control may be that desirability judgments vary among sample populations, as some authors have reported.[21,22]

Jackson[23] aimed for control of SD response set in the construction and selection of items for the Personality Research Form. Yet a study by Braun and Asta[24] showed that instructions to fake and make an excellent impression produced significant changes in all but two of the inventory's scales. With instructions to fake a bad score, Hoffmann[25] discovered significant changes in fourteen of twenty-two scales. Such findings may not be proof that scores are inaccurate under

---

20 John R. Braun, "Forced-Choice Self-Report Devices: A Look at Some Unwarranted Claims," Measurement and Evaluation in Guidance, Vol. 2, No. 3, 1969, p. 153-156.

21 Nancy Wiggins, "Individual Viewpoints of Social Desirability," Psychological Bulletin, Vol. 66, No. 2, 1966, p. 68-77.

22 Donald W. Fiske and Pamela H. Pearson, "Theory and Techniques of Personality Measurement," Annual Review of Psychology, Vol. 21, 1970, p. 72.

23 Jackson, Personality Research Form Manual, p. 15.

24 John R. Braun and Patricia Asta, "Changes in Personality Research Form Scores (PRF, Form A) Produced by Faking Instructions," Journal of Clinical Psychology, Vol. 25, No. 4, 1969, p. 429-430.

25 Helmut Hoffmann, "Performance on the Personality Research Form Under Desirable and Undesirable Instructions: Personality Disorders," Psychological Reports, Vol. 23, No. 2, 1968, p. 507-510.

standard instructions, but they do illustrate that subjects are able to identify "good" responses and "bad" responses to items in which social desirability was meant to be minimized.

Reviews of the problem of response styles in personality measurement illustrate that, despite the abundance of research data, there is often disagreement on the definition of essential terms and on techniques of measuring response styles. As well, research evidence is equivocal whether response styles are traits which can be generalized and which are predictive of other non-test behavior.[26] In their review of personality measurement Fiske and Pearson described the measurement of response sets as chaotic, and they concluded: "The majority view concerning the role of social desirability and acquiescence tends to be that such sets do in fact exist, but do not appear to be as pervasive as once supposed."[27]

The existence and influence of social desirability response set remained in dispute in the literature. Yet there was sufficient evidence which supported SD response set as a confounding source of variance that improvements in self-report personality inventory were warranted. The error-choice technique was proposed as a method of avoiding the influence

---

26 McGee, "Response Set in Relation to Personality: An Orientation," p. 27-28.

27 Fiske and Pearson, "Theory and Techniques of Personality Measurement," p. 49-86.

of SD response set. That claim, and a rationale hypothesized for the technique, merited experimental investigation.

2. The Error-Choice Technique.

The error-choice technique was originally devised as an indirect method of assessing attitudes with accuracy and subtlety. Hammond,[28] noting earlier research that attitudes are capable of producing systematic error in perception and recall, proposed a technique of attitude survey which would measure the constancy and direction of the error. He theorized that since the field of social events and personalities presents complex and confusing stimuli to the observer one could expect the perception of such stimuli to be distorted by social or attitudinal factors. Similarly, if a person were remembering a previous perception, one could expect his recall to be selective, fitting his frame of reference or supporting his established premise.

To investigate the hypothesis, surveys of attitudes about Russia and about labor management were conducted by means of what Hammond called the error-choice[29] technique. A questionnaire in multiple choice format was presented to

---

28 Kenneth R. Hammond, "Measuring Attitudes by Error-Choice: An Indirect Method," Journal of Abnormal and Social Psychology, Vol. 43, No. 1, 1948, p. 38-49.

29 The abbreviation EC will be used for "error-choice."

subjects ostensibly as a test of information. For each item there were two answer choices, of which neither was correct, so that the subject was forced into a choice of errors. The "error answer choices" were determined either by putting the answers equidistant from the truth in opposite directions, or by using questions where the truth was indeterminable and putting the answers to opposite extremes. An example of the former type of item follows: "Average weekly wage of the war worker in 1945 was (1) $37, (2) $57." The average wage was actually $47.

The results from the questionnaires bore out the theory that attitudes would distort in a consistent direction either the original perception or the subsequent recall of facts. Hammond's subjects were union employees and businessmen, and he considered them criterion groups for opposing attitudes. The groups produced constant errors opposite in direction, statistically significant for both topics of the survey.

A corollary study showed that responses might differ for an item depending upon whether it was presented as an attitude survey item or an information test item. On the basis of these results, Hammond concluded that the EC technique could be used to eliminate response set from attitude surveys since respondents revealed their attitudes

while being unaware of the implications of their error choices when items were presented as a test of information.

Two similar studies distinguished subjects' attitudes under the guise of testing their knowledge of facts. Weschler[30] developed an inventory in EC format to assess attitudes toward labor and management. He divided subjects into two groups on the basis of their supposedly anonymous report of sympathy with labor or with management. Subjects were administered the EC inventory and were instructed to guess intelligently when they did not know the answer. Raw scores from the inventory differentiated the groups significantly ($p < .01$). When items were weighted according to their critical ratios, the weighted inventory scores differentiated better although there was still a very considerable overlap between the groups. With only eleven EC items in the inventory, the test-retest reliability was .62.

Kubany[31] investigated the ability of an EC attitude questionnaire about socialized medicine to differentiate between two groups of "known bias," medical and social work students respectively. EC non-factual items were interspersed among legitimate factual items. A significant difference

---

30 Irving R. Weschler, "An Investigation of Attitudes Toward Labor and Management by Means of the Error-Choice Method: I," Journal of Social Psychology, Vol. 32, 1950, p. 51-62.

31 Albert J. Kubany, "A Validation Study of the Error-Choice Technique Using Attitudes on National Health Insurance," Educational and Psychological Measurement, Vol. 13, No. 2, 1953, p. 157-163.

between the known groups was reported. However, this difference appeared only after some items were discarded from the calculation because their validity coefficients were too low. Odd-even reliability of EC items was .90, corrected by the Spearman Brown formula. No significant relation ($r = -.03$) was found between attitude, as measured by direction of the error-choice, and the amount of factual knowledge about the topic. The investigator concluded that the EC format was useful to discover the subjects' "private" attitudes.

Years after the original investigations of the theory, Wilde and Fortuin[32] proposed adapting the EC technique to personality assessment as a possible remedy for the fakeability and the contaminating variance from response sets which were ascribed to personality inventories. They transformed Wilde's Dutch personality questionnaire[33] into EC format. They first administered the inventory to a rather heterogeneous sample population in order to obtain the endorsement frequency for each item; that is, the proportion of the sample answering YES to the item. The two answer choices for the EC version were prepared by adding a small proportion to the endorsement frequency and subtracting the

---

32 G.J.S. Wilde and S. Fortuin, "Self-Report and Error-Choice: An Application of the Error-Choice Principle to the Construction of Personality Test Items," British Journal of Psychology, Vol. 60, No. 1, 1969, p. 101-108.

33 G.J.S. Wilde, Neurotische Labiliteit gemeten volgens de Vragenlijstmethode (The Questionnaire Measurement of Neurotic Instability), Amsterdam, Van Rossen, 1963.

same proportion from the endorsement frequency. Thus, the item "Do you often suffer from headaches?" had an endorsement frequency of .21 in the sample population. The EC version of the item might then be: "Do you often suffer from headaches? (a) 26% (b) 16%." The EC version was presented to subjects as a test of their insight into people, where the task was to judge how others had answered the items. Subjects were told that a large sample of people answered the questionnaire, allowing the percentage of people answering YES to be calculated for each item. Of the two percentages following each item, subjects were instructed to choose the correct one; that is, the percentage of people who did answer YES.

Wilde and Fortuin[34] hypothesized a relationship between subjects' error-choice responses and their self-report[35] responses to the same items. They theorized that choosing the higher percentage in error-choice format (i.e., overestimating the endorsement frequency of the population) would correspond to self-report endorsement; and, conversely, that EC underestimates would correspond to self-report NO responses. Such a correspondence between self-report and error-choice answers would allow inferences to be made about the subject's personality from his responses to the EC questionnaire.

---

34 Wilde and Fortuin, "Self-Report and Error-Choice: An Application of the Error-Choice Principle to the Construction of Personality Test Items," p. 103.

35 The abbreviation SR will be used for "self-report."

It was hypothesized that confounding variance from SD response set would be eliminated from the error-choice inventory. For that reason, correspondence between self-report and error-choice responses should not be perfect. SD response set, operating in self-report inventory, would reduce to some degree the accuracy of the trait scores as measures of a subject's personality. If SD response set were not a contaminating variable in EC inventory, then trait scores derived from EC inventory would be different from self-report scores; they might be more accurate measures of a subject's traits.

Wilde and Fortuin obtained some empirical support for their hypothesis. They administered Wilde's personality inventory to 150 subjects, in both self-report and error-choice versions. Ninety of the ninety-one items showed a positive relation between affirmative SR responses and EC overestimations. In sixty-two items, X^2 was significant at the .01 level, and in an additional eleven items at the .05 level. One item showed a negative relation that was insignificant. The SR-EC correspondence was found in 68.24 per cent of all items, ranging from 73.98 per cent in items of the Somatic Neuroticism scale to 62.01 per cent in items of the Extroversion scale.[36]

---

36 Wilde and Fortuin, "Self-Report and Error-Choice: An Application of the Error-Choice Principle to the Construction of Personality Test Items," p. 105.

In another paper reporting the same study, Wilde and de Wit[37] concluded that the SR-EC relation existed for most subjects. Their conclusion was based on a liberal criterion: for 141 of the 150 subjects the relation held in more than fifty per cent of the items.

Wilde and de Wit also investigated the validity of personality trait scores derived from EC inventory. Scores on the inventory's scales were derived for subjects by counting one point for each EC overestimation when the corresponding SR item was YES-keyed, and one point for each EC underestimation when the SR counterpart was NO-keyed. That would be the method of deriving a testee's scores for the inventory's traits if the EC format were used as an indirect assessment of the testee's personality. These EC trait scores were related to their respective SR trait scores, correlations ranging from .17 to .54.[38] The gross criterion measure of personality was as follows: subjects who were being treated for neurotic conditions were considered poorly adjusted; all other subjects were considered well adjusted. Self-report trait scores had correlations with the criterion ranging from .06 to .60 in absolute values. Except for the

---

37 G.J.S. Wilde and O.E. de Wit, "Self-Report and Error-Choice: Inter-Individual Differences in the Operation of the Error-Choice Principle and Their Validity in Personality Questionnaire Tests," British Journal of Psychology, Vol. 61, No. 2, 1970, p. 222, 224.

38 Ibid.

Lie scale, EC trait scores had a relation to the criterion that was in the same direction as the SR scores; the correlations for EC scores were lower than those for the respective SR scores. The investigators suggested the validity coefficients for the EC trait scores were conservative because the EC scales were derived directly from the SR scales without sharpening through item analysis and because the validity coefficients were not corrected for attenuation.

Wilde and his colleagues were encouraged sufficiently by their data that they considered the error-choice format of personality inventory worthy of further study. Its promise was that it might eliminate from trait scores the construct-irrelevant variance contributed by social desirability response set. The present thesis aimed to investigate the influence of SD response set in the correspondence between SR and EC inventories. The experimental hypothesis is developed in the next section.

3. SD Response Set and Error-Choice Inventory.

Wilde and Fortuin[39] hypothesized a cognitive process that might underlie the correspondence between self-report and error-choice responses. It was predicated on the

---

39 Wilde and Fortuin, "Self-Report and Error-Choice: An Application of the Error-Choice Principle to the Construction of Personality Test Items," p. 103.

influence of SD response set and the need for dissonance reduction in subjects. The rationale was as follows.

For every item in self-report, the subject is motivated to answer in the desirable direction; i.e., to endorse desirable items and not to endorse socially undesirable items. But there is an opposing motivation as well, either honesty, when the subject's honest opinion differs from the SD response; or modesty, when the subject's honest opinion coincides with the SD response. (The widespread existence of modesty as a motivation in responding seemed a tenuous element in the theory.) Thus for every item the subject experiences dissonance between SD motivation and either honesty or modesty motivation.

In desirable items, SD promotes a YES response. If the subject answers YES, he has overcome the opposing influence of honesty or modesty perhaps by rationalizing that "most people are that way," or that "most people would endorse the item." (The error-choice questionnaire requires subjects to judge the frequency of endorsement; some subjects may make a distinction between people's actual behaviors and their reported behaviors, others may equate the two.) The rationalization that most people would endorse the item would serve to reduce the dissonance within the subject, perhaps by increasing the strength of the SD motivation, or perhaps by adding to one side of the dissonance equation a motivation to be like other

people. (Wilde and Fortuin did not specify how the rationalization would operate to reduce dissonance.) Thus, by adjusting his cognitions about other people, the subject makes others like himself, conceives them in his own image. When faced with the same item in EC format, the subject maintains his rationalization that "most people are that way" and chooses the higher of the two percentages available as response alternatives. He chooses the EC overestimation.

If the subject answers YES to an undesirable item, honesty has overcome the opposing SD motivation. The subject has resisted SD motivation, according to the theory, by again rationalizing that "most people are that way." He thereby reduces the dissonance, perhaps by adding to the strength of the honesty motivation. When faced with the same item in EC format the subject chooses the overestimation. Thus the rationale provides that whenever the subject endorses an item, whether it be in the desirable or the undesirable direction, a process of dissonance reduction leads him to choose the EC overestimation. That accounts for correspondence between self-report endorsement and EC overestimation.

Similarly, whenever the subject answers NO to an item, he has had to resolve a conflict between SD motivation and honesty or modesty motivation. It was hypothesized he does so by rationalizing that "most people are not that way," or that they would not endorse the item. Therefore, when

faced with these items in EC format, he chooses the lower of the two percentages available. So, the SR endorsement-EC overestimation relation, in inverted form, holds: self-report NO responses correspond to EC underestimations.

The present investigator questioned whether a more parsimonious rationale was not more likely than that offered by Wilde and his colleagues. Did SD response set underlie both self-report endorsement and EC overestimation, thus accounting for the correlation between them? If that were so, trait scores derived by EC inventory would be no better measures of personality dimensions than self-report scores; they would suffer from the same confounding variance that was claimed to affect SR scores, and for which EC format was proposed as a remedy.

There was some evidence that social desirability response set does operate when subjects are asked to rate the frequency of endorsement of items. By means of a multimethod factor analysis, Kusyszyn and Jackson[40] identified a desirability factor that existed in subjects' judgments of endorsement frequency as well as in their endorsement of items. Another study also suggested that desirability was a factor in estimating endorsement frequencies. Jackson and

---

40 Igor Kusyszyn and Douglas N. Jackson, "A Multimethod Factor Analytic Appraisal of Endorsement and Judgment Methods in Personality Assessment," Educational and Psychological Measurement, Vol. 28, No. 4, 1968, p. 1047-1061.

Messick[41] noted that subjects appeared to recognize the existence of social desirability response bias in people's endorsement of inventory items. When subjects judged the frequency of endorsement, their ratings had a higher correlation with item SDSVs than did their judgments of the frequency of occurrence of item content. Subjects' estimates of item endorsement correlated .87 with SDSVs; their estimates of occurrence of item content correlated .71 with SDSVs.

Some of the data reported by Wilde and his co-authors could have suggested the influence of SD response set in error-choice responding. Wilde and de Wit correlated the strength of SR-EC correspondence shown by subjects with their self-report scores on the scales of the inventory. The strength of the SR-EC relation had a correlation of .31 with scores on the Social Desirability scale, and negative correlations with scores on Neuroticism ($r = -.38$) and Somatic Neuroticism ($r = -.29$) scales.[42] The two neuroticism scales could be expected to be composed of items of socially undesirable content. These three correlations could be evidence of

---

41 Douglas N. Jackson and Samuel Messick, "A Distinction Between Judgments of Frequency and of Desirability as Determinants of Response," Educational and Psychological Measurement, Vol. 29, No. 2, 1969, p. 273-293.

42 Wilde and de Wit, "Self-Report and Error-Choice: Inter-Individual Differences in the Operation of the Error-Choice Principle and Their Validity in Personality Questionnaire Tests," p. 222.

a consistent, moderate relation between strength of SR-EC correspondence and SD response set. Wilde and his colleagues, however, did not make such an interpretation.

Wilde and de Wit also performed a centroid factor analysis on all variables (strength of SR-EC correspondence for each inventory scale, average SR scores for each scale, average EC scores for each scale, and criterion of adjustment) with unity in the diagonal cells. Of the four factors extracted, factors I and II were rotated so that factor I[1] would pass through the variable, criterion of adjustment. Highest loadings on this factor came from self-report scores on Neuroticism, Somatic Neuroticism, and Social Desirability (-.78, -.78, .80 respectively). Error-choice scores on these same scales loaded almost as highly: -.76, -.71, .75. The factor was interpreted as an adjustment factor.[43] But perhaps it might be interpreted as SD response set, operating in both self-report and error-choice versions. As Edwards[44] noted, individual differences in rates of SD responding are very similar to, and highly correlated with, individual differences in rates of "normal, healthy and adjustive" responding.

---

43 Ibid., p. 225-226.

44 Edwards, The Measurement of Personality Traits by Scales and Inventories, p. 232.

Thus there was evidence from the studies by Wilde and his colleagues and also from other sources that SD influence was not absent from judgments of endorsement frequencies. It was not certain that the error-choice inventory eliminated variance from SD response set, as its advocates suggested. Moreover, the rationale hypothesized by Wilde to account for self-report-error-choice correspondence was rather extravagant and quite unproven. There was need to investigate whether SD response set accounted for the SR-EC correspondence.

Another hypothesis for experimental testing was developed from a second rationale proposed by Wilde and co-authors. The relevant theory is the subject matter of the following section.

4. Conformity and Error-Choice Inventory.

Wilde and Fortuin[45] also offered an alternative to the dissonance reduction process as a rationale for the SR-EC correspondence. The alternative rationale assumed a conformity motivation in subjects, as follows. Subjects are motivated to describe themselves in terms of the norm. Their responses to self-report personality inventories depend upon

---

45 Wilde and Fortuin, "Self-Report and Error-Choice: An Application of the Error-Choice Principle to the Construction of Personality Test Items," p. 103.

their estimation of how the majority of people would answer. Subjects then make their self-descriptions conform to what they estimate most people are like; or they may conform to what they estimate most people would report they are like.

According to the rationale, when subjects are presented with the error-choice version of an inventory they merely record the estimate of the population that was implicit during their self-report responding. If a subject had estimated most people would endorse an item, he would choose the larger of the two percentages available as answer choices; if he had estimated most people would not endorse an item, he would choose the smaller EC percentage. Thus the rationale provides that whenever a subject endorses an item in self-report he chooses the error-choice overestimation, and that SR NO responses correspond to EC underestimations.

Wilde and his colleagues apparently attempted no investigation of the hypothesized process of conforming to the population norm. From the description of the process, it seemed that a personality dimension of conformity-independence[46] should be related to the amount of correspondence between self-report and error-choice responding. Highly conforming subjects could be expected to show a large amount of SR-EC correspondence, while highly independent subjects should

---

46 "Conformity," "conformity-independence" and "conformity-proneness" will be used interchangeably as equivalent terms for the personality dimension.

show little correspondence. If that were so, error-choice responses of independent subjects would not allow correct inferences to be made about the subjects' personalities. That possible limitation of the error-choice format of personality inventory needed to be investigated.

The existence and measurement of conformity as a personality trait were topics of some dispute in the literature. However, the definition by Krech, Crutchfield and Ballachey[47] was supported by experimental data, and it appeared applicable to Wilde and Fortuin's rationale. Krech et al. conceptualized conformity as essentially the yielding to group pressures when there is conflict between forces within the individual which "tend to lead him to act, value and believe in one way and those pressures emanating from the society or group which tend to lead him in another way." The group pressure need not be overt: group pressure exists whenever there is a group judgment or action which differs from the individual's own. On the basis of laboratory study, Krech et al. argued for the existence of "stable and enduring conformity tendencies in people--in short, for an interpersonal response trait of conformity-proneness."[48]

---

[47] David Krech, Richard S. Crutchfield and Egerton L. Ballachey, Individual in Society, New York, McGraw-Hill, 1962, p. 506.

[48] Ibid., p. 523.

Crutchfield[49] used a laboratory procedure to measure the trait in subjects. On the basis of measures from personality inventories and projective techniques he found that conformists tend, among other things, to exhibit intense preoccupation with other people, as contrasted with the more self-contained, autonomous attitudes of the independent persons; to be less insightful and realistic in their perception of themselves; to be poorer in ability to judge other people's traits accurately; to be more passive and suggestible in their interpersonal behavior.

In further investigations, subjects were observed and interviewed over a three-day assessment period by "expert psychological observers." Some of the Q-sort descriptions applied to highly independent subjects were "self-reliant; independent in judgment; able to think for himself." Some of the descriptions applied to highly conforming subjects were "lacks insight into his own motives and behavior"; "judges self and others in conventional terms, such as 'popularity,' the correct thing to do, group opinion, etc."

Those findings suggested that those who are highly conforming in the Crutchfield laboratory are very likely to be conforming, and to lack individuality and independence of thought and action, in other situations. One might expect the conformists to lack a clearly differentiated

---

49 Ibid., p. 525-528.

perception of self and perception of others; the conformists might keep their behavior or, at least, their self-descriptions in line with some accepted norm. For either of these reasons, one could expect a correspondence between conformists' self-descriptions and their descriptions of others; that is, for conformists, self-report answering of inventories should be similar to their judgment of others in error-choice inventory. On the other hand, one could expect the independents in the Crutchfield situation to have a differentiated perception of self and perception of others, to show individuality in their behavior and self-descriptions. Thus, their self-report answering of inventories should not correspond to their judgment of others in error-choice inventory.

The present thesis investigated whether conformity, as measured by the Crutchfield laboratory procedure, was a factor in the amount of SR-EC correspondence shown by subjects.

The rationale based on a conformity motivation was one of two offered by Wilde and his colleagues. It hypothesized that a subject's estimate of other people influenced his self-report responses. In contrast, Wilde's alternative rationale of dissonance reduction hypothesized that a subject's self-report responses affected his estimate of other people; that is, self-report responses influenced EC responses.

Wilde allowed that either rationale might account for SR-EC correspondence, or that both hypothesized processes might occur together. In their study of self-report and error-choice responding, Wilde and Fortuin administered the EC version after the SR inventory. Since the two rationales presented opposing theories as to which inventory influenced the other, there was need to investigate the order of administration of the inventories as a factor in SR-EC correspondence, either alone or in interaction with the possible factors of conformity and SD response set.

5. Summary and Experimental Hypotheses.

The review of the literature revealed that the existence and the influence of social desirability response set are in dispute. However, the evidence for such a response set in self-report inventory had gained qualified acceptance among a wide audience. Consequently there continued to be a need for new methods of personality measurement which would avoid contamination of trait scores by socially desirable responding.

The error-choice technique was demonstrated to have some value as an indirect method for assessing attitudes while subjects remained unaware that their attitudes were being examined. When the technique was adapted for use with personality inventory, Wilde and his colleagues concluded

that it showed promise as a means of eliminating the influence of SD response set. They hypothesized two alternative rationales that might underlie a correspondence between error-choice and self-report responses.

One rationale was predicated on the influence of social desirability in self-report but not in error-choice inventories. The present investigator hypothesized that SD response set might operate in EC responding as well as self-report responding, thereby accounting for SR-EC correspondence. The present thesis aimed to investigate the relation of SD response set in subjects to the amount of SR-EC correspondence they showed. The experimental hypothesis reads as follows:

> The degree of SD response set in subjects is a significant factor in the amount of correspondence between their self-report and error-choice responses.

The second rationale hypothesized by Wilde and his colleagues to account for SR-EC correspondence depended on a conformity motivation in subjects. The present thesis aimed to investigate the relation of conformity-proneness, as measured by the Crutchfield laboratory procedure, to the amount of SR-EC correspondence shown by subjects. The experimental hypothesis reads as follows:

> The degree of conformity-proneness in subjects is a significant factor in the amount of correspondence between their self-report and error-choice responses.

The two rationales hypothesized by the proponents of error-choice personality inventory offered contradictory explanations. One held that EC responses influenced SR responses, the other that SR influenced EC. For that reason the present investigator considered it important to evaluate the order of administration of the self-report and error-choice inventories as a factor in SR-EC correspondence. The pertinent experimental hypothesis reads as follows:

> The order of administration of self-report and error-choice inventories is a significant factor in the amount of correspondence in subjects' responses.

Finally, the present thesis investigated for interaction among the three factors noted in the preceding experimental hypotheses. Thus, the last experimental hypothesis reads:

> One or more interactions of SD response set, conformity-proneness, and order of administration significantly affect the amount of correspondence between subjects' self-report and error-choice responses.

CHAPTER II

EXPERIMENTAL DESIGN

This chapter reports details of the method adopted in order to test experimentally the working hypotheses presented at the end of Chapter I. First, the sample of subjects is described. The design and general procedure of the experiment are next elaborated. In the following section the tools are described, and also the construction of the error-choice version of the personality inventory. Then the statistical operations applied to the data are noted, and finally the hypotheses of the experiment are presented in the null form.

For the experiment two measures were obtained for each subject, a behavioral measure of conformity and an inventory measure of social desirability response set. On the basis of both scores subjects were trichotomized. Then a self-report inventory and an error-choice inventory were administered to subjects, half of the subjects in each group receiving the inventories in one order and half in the reverse order. Scores were obtained which indicated the amount of SR-EC correspondence for each subject, and they were submitted to statistical analysis.

1. The Sample.

Subjects were 108 female twelfth grade graduates between the ages of seventeen and twenty-four inclusive. Sixty-two were students at a hospital school of nursing, thirty-six were students in a secretarial science program at an institute of technology, and ten were university students. All were volunteers.

2. The General Method of the Experiment.

One hundred and ten subjects were listed alphabetically and the list was divided from beginning to end into groups of five, since the Crutchfield procedure necessitates that five subjects be assessed at the one time. The groups were randomly assigned to receive either the conformity measure first or the measure of SD response set first. Subjects were then scheduled for testing.

Upon arrival at the laboratory, subjects were met by Experimenter 1 and were introduced to Experimenter 2, who assisted in the Crutchfield procedure. By means of first-name introductions and the expression of appreciation for subjects' participation, the experimenters' manner was intended to be non-authoritarian. Either the conformity measure or the scale measuring SD set was administered to the five subjects, according to the randomly determined order of presentation.

When the SD scale was administered first, the conformity measure was administered immediately after. When the latter was completed, Experimenter 1 explained the deception involved in the conformity measure: he mentioned the findings of other investigators, he interpreted the conforming behavior in a positive way, and he expressed the hope that subjects felt the deception was justified. Subjects were urged not to discuss the nature of the deception with classmates who might later participate in the experiment. When the conformity measure was administered first to a group, the explanations were offered immediately afterward--before the administration of the SD scale.

Instructions for the SD scale were printed at the beginning of the inventory and were read aloud by Experimenter 1. The introduction and general instructions for the conformity measure also were read by him; the separate statement of the task for each item was read by Experimenter 2.

After all groups had been administered the two measures, a random discarding procedure was used to reduce the number of subjects to the multiple of eighteen required by the design of the experiment; here, 108 subjects. The correlation was calculated between conformity scores and SD response set scores to ascertain that they were independent measures.

Subjects were then trichotomized into low, average, and high groups on the basis of SD scores. Each third of the sample was then trichotomized on the basis of conformity scores, giving nine groups of twelve subjects each. Half the subjects within each of the nine groups were randomly assigned to answer the SR inventory first, half to answer the EC inventory first. The procedure of randomizing the treatment factor for all levels of the classification factors was outlined by Edwards.[1]

Within four weeks of the first testing session, subjects, in groups of varying size, completed the SR and EC inventories. The second inventory was administered as soon as the first one was completed by all subjects who were present. Instructions to both inventories were read aloud by the experimenter. Subjects were told to answer every item in both inventories.

For each subject an SR-EC correspondence score was derived: the number of times that a self-report item was endorsed and the counterpart error-choice item was answered with the overestimate, plus the number of times that a self-report item was answered FALSE and the counterpart error-choice item was answered with the underestimate.

---

1 Allen L. Edwards, Experimental Design in Psychological Research, New York, Holt, Rinehart and Winston, 1968, p. 260-262.

Statistical analysis of the SR-EC correspondence scores was performed to investigate the influence of SD response set, conformity, and order of administration.

3. The Tools and Related Procedures.

This section describes the measure of SD response set and the behavioral measure of conformity which were administered to each subject. Then the self-report personality inventory is identified, and the method for constructing the error-choice version is reported.

A. The Measure of SD Response Set

A scale constructed by Edwards,[2] using items from the MMPI, was employed as the measure of SD response set. Edwards had ten judges rate 150 items for social desirability. Of the seventy-nine items that were rated unanimously, item analysis selected the best thirty-nine to form the SD scale. If an item had a socially desirable scale value, a YES response to it was an SD response; if an item was socially undesirable, a NO response was keyed for SD. Edwards' SD scale was "designed to measure the tendency of subjects to give socially desirable responses under the usual or

---

2 Allen L. Edwards, The Social Desirability Variable in Personality Assessment and Research, New York, Dryden, 1957, p. 30.

standard instructions of administration of personality scales."[3] A reliability coefficient of .83 was reported.[4]

If the SD measure was to be a variable in a factorial experiment, it had to be independent of any other classification or organismic factor. Since conformity-proneness was to be a classification factor, the SD measure had to be uncorrelated with conformity as measured by the Crutchfield technique.

The relationship between conformity and SD measures was not clearly defined in the literature. As noted previously, the highly conforming individual in the Crutchfield situation was described as "judges self and others in conventional terms, such as 'popularity,' the correct thing to do, group opinion, etc." The description suggested that conformists would be more susceptible to SD influence than would independents when answering self-report inventories. However, Back and Davis[5] found no simple correlation between conformity and a measure of socially desirable responding.

---

3 Allen L. Edwards, Carol J. Diers and Jerald N. Walker, "Response Sets and Factor Loadings on Sixty-One Personality Scales," Journal of Applied Psychology, Vol. 46, No. 3, 1962, p. 224.

4 Allen L. Edwards and James A. Walsh, "Response Sets in Standard and Experimental Personality Scales," American Educational Research Journal, Vol. 1, No. 1, 1964, p. 56.

5 Kurt W. Back and Keith E. Davis, "Some Personal and Situational Factors Relevant to the Consistency and Prediction of Conforming Behavior," Sociometry, Vol. 28. No. 3, 1965, p. 227-240.

(The latter was a scale not in common use and having a reliability coefficient of only .59.) The SD measure was related to conformity in the Crutchfield situation only for subjects who considered themselves a "close friend" of another member of the experimental group.

There was some indirect evidence that Edwards' SD scale was not related to the Crutchfield measure of conformity. In a study of the concept of social desirability, Marlowe and Crowne[6] administered Edwards' SD scale and Barron's "Independence of Judgment Scale." (Barron[7] had measured subjects in an Asch conformity situation; he selected twenty-two inventory items to form a scale that was highly related to the yielding type of conformity.) Marlowe and Crowne found no significant relation (r = -.12) between Edwards' SD scale and conformity as measured by Barron's scale. Nor was Edwards' SD scale related to an experimental measure of tendency to express socially appropriate opinions. Thus Edwards' SD measure appeared not equivalent to a personality dimension of conformity-proneness.

---

6 David Marlowe and Douglas P. Crowne, "Social Desirability and Response to Perceived Situational Demands," Journal of Consulting Psychology, Vol. 25, No. 2, 1961, p. 109-115.

7 Frank Barron, "Some Personality Correlates of Independence of Judgment," Journal of Personality, Vol. 21, No. 3, 1953, p. 287-297.

A study by Tuddenham[8] found that conformity in the Crutchfield situation was not significantly related to a scale measuring a subject's interest in creating a good impression, the Good Impression scale of the California Psychological Inventory.[9] After finding few significant correlations between the Crutchfield measure and various inventory trait scales, Tuddenham concluded that for women conformity in the Crutchfield situation is not predictable from personality trait measures. That conclusion suggested that conformity would not be related to an SD scale.

There remained the need to verify the independence of the conformity and SD response set measures. It was to be the first statistical analysis before proceeding with the experimental design originally proposed.

The thirty-nine items of Edwards' SD scale were administered as a separate self-report inventory identified as "Introductory Inventory."[10] The range of scores possible was from zero to thirty-nine.

---

8 Read D. Tuddenham, "Correlates of Yielding to a Distorted Group Norm," Journal of Personality, Vol. 27, No. 2, 1959, p. 272-284.

9 Harrison G. Gough, California Psychological Inventory, Palo Alto, Consulting Psychologists, 1956, 12 p.

10 A copy is included in the appendix.

B. The Measure of Conformity

Subjects were measured on a personality dimension of conformity-independence by means of a laboratory procedure designed to emulate Asch's[11] situation of group influence. The procedure was developed by Crutchfield[12] and modified by Tuddenham et al.[13] The laboratory is illustrated in Figure 1. The apparatus and procedure were as follows.

The laboratory was 23 feet in length and 13 feet 9 inches in width. At one end was a projection screen which offered a surface 60 inches square. Toward the opposite end of the room were five booths, formed by opaque side panels but open at the front and back. The panels were 82 inches high and they defined the booth area as 26 inches in width and 45 inches in depth. The booths were arranged in a row so that their fronts gave a view of the projection screen, 15 feet opposite. Each booth contained a chair and a small table on which rested an answering console. A slide projector with a 4 inch lens was positioned on the table in the middle booth, at the back of the answering console. When

---

11 S.E. Asch, "Effects of Group Pressure Upon the Modification and Distortion of Judgments," in Harold Guetzkow (ed.), Groups, Leadership and Men, Pittsburgh, Carnegie, 1951, p. 177-190.

12 Richard S. Crutchfield, "Conformity and Character," American Psychologist, Vol. 10, No. 5, 1955, p. 191-198.

13 Read D. Tuddenham, Philip Macbride and Victor Zahn, Studies in Conformity and Yielding I, Technical Report No. 1, Contract NR 170-159, Berkeley, University of California, 1956, iii-72 p.

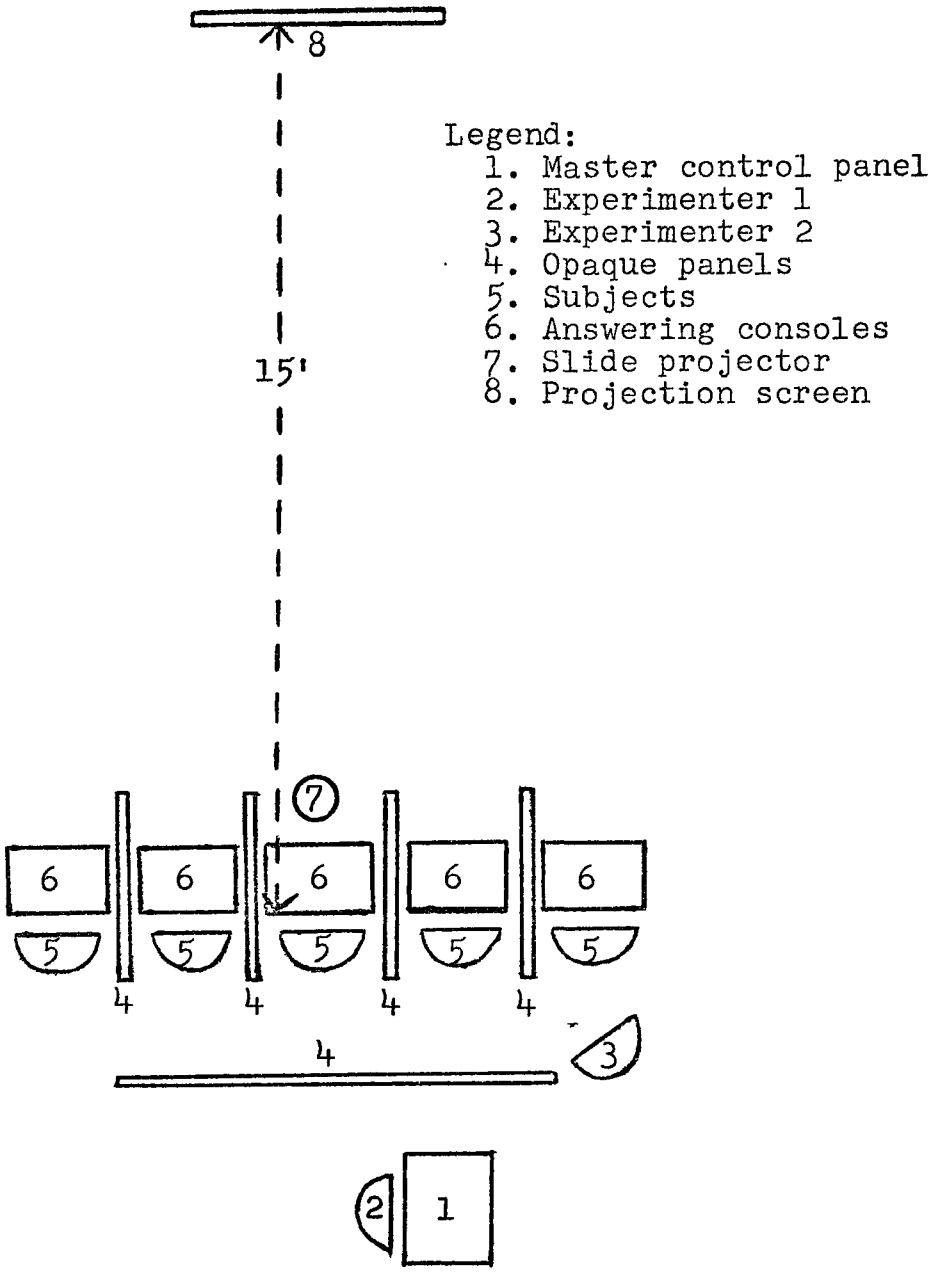


Figure 1.- Diagram of the Laboratory for Measuring Conformity-Proneness.

Adapted from Read D. Tuddenham et al., Studies in Conformity and Yielding I, Technical Report No. 1, Contract NR 170-159, Berkeley, University of California, 1956, iii-72 p.

subjects were seated in the booths they could not see one another, nor one another's consoles.

An opaque panel 7 feet in width and 82 inches in height was located approximately 3 feet behind the row of booths. This panel hid from the subjects' view the master control panel, where Experimenter 1 sat. Experimenter 2, who operated the projector, sat behind one of the end booths.

The answering console in each booth had a facade measuring 19 by 10-1/2 inches. The console comprised five horizontal rows of signal lights and below them a row of nine silently operating toggle switches. The switches were numbered from one to nine, and the signal lights were arranged in columns above the switches so that within each of the five rows there was a light which corresponded to each of the switches. The signal lights were green, 5/8 inch in diameter, and spaced 1-3/8 inches apart horizontally. The rows of lights were spaced 1-1/2 inches apart and were marked by a red signal light to the left of each row, separated by extra spacing from the green lights. The red lights were numbered one to five. The console is illustrated in Figure 2.

There were two modes of operation for the electronic equipment. Subjects were given to understand that the apparatus operated in the following mode only, the automatic mode. By means of the master control, Experimenter 1

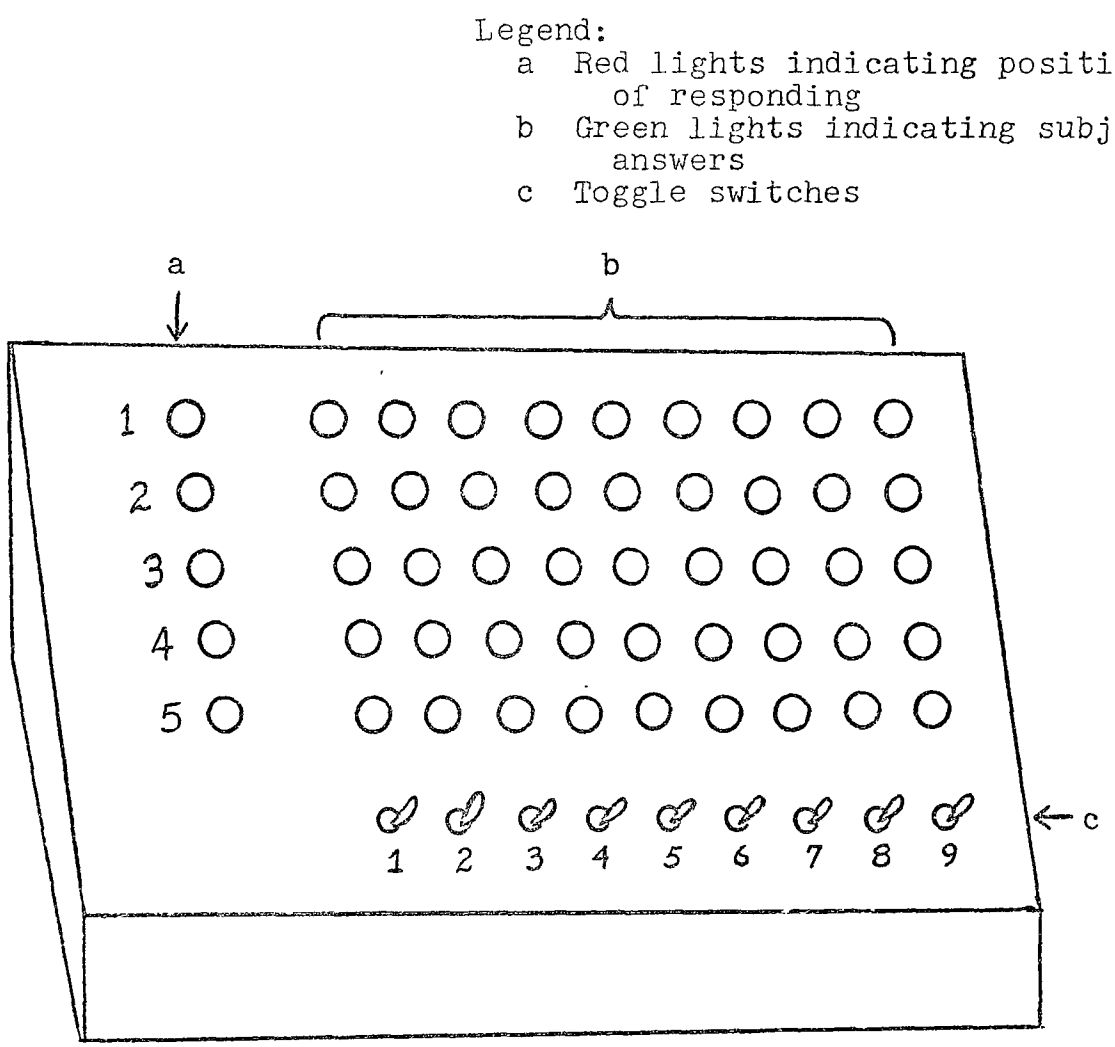


Figure 2.- Diagram of the Answering Console Used in the Procedure for Measuring Conformity-Proneness.

Adapted from Read D. Tuddenham et al., Studies in Conformity and Yielding I, Technical Report No. 1, Contract NR 170-159, Berkeley, University of California, 1956, iii-72 p.

assigned to each subject her position of responding--first, second, third, fourth or fifth. The order of responding was communicated to subjects by means of the numbered red lights on the consoles. If the red light opposite row three was illuminated on a subject's console, for example, that subject was to respond third. She was to wait until subjects who were first and second in order had responded before she answered. Subjects were instructed that they must answer in order and that their position of responding would be changed on every trial.[14]

The nine toggle switches on the consoles represented answer choices one to nine. If a subject was designated to answer third, for example, when she activated one of the switches the corresponding green light would be illuminated in row three on her console and also in row three on the consoles of the four other subjects. Similarly, when the subject designated to answer first activated a switch, the green light corresponding to the switch would be illuminated in row one on all consoles. Thus the green lights indicated to all subjects <u>when</u> the subject first in order had answered, <u>when</u> the subject second in order had answered, and so on. In addition, the green lights revealed to each subject <u>which</u>

---

[14] The instructions to the subjects are included in the appendix.

answers were chosen by the others. A subject's answer also registered on the master control panel.

The second mode of operation for the equipment, the manual mode, was not revealed to the subjects. In this mode a subject's answer registered only on her console and on the master control panel, but not on the consoles of the four other subjects. From the master control, Experimenter 1 could simulate answers of the four other subjects on each of the consoles. For a trial, all subjects were assigned the same position of answering, although they were under the impression that the five positions were distributed among them. When they were all designated first to answer, for example, Experimenter 1 waited for every subject to respond and then simulated the second, third, fourth and fifth answers. When subjects were all designated fifth to answer, they waited their turn while Experimenter 1 simulated the answers of four subjects; then each subject answered, believing herself to be fifth to do so.

For each group of five subjects there was an instruction period during which they had three practice trials with the equipment operating in the automatic mode. The equipment was changed to the manual mode for the rest of the session without the subjects' knowledge.

The task for subjects was presented by means of projected slides of 35 millimeter format, together with a

verbal statement of the task by Experimenter 2. The series of items was constructed by Tuddenham[15] and posed problems of visual discrimination, such as comparisons of line length, slope or convergence, and comparisons of areas, proportions and trajectories. Items were designed to be simple enough for children aged ten to twelve years, and critical items had to have a consensus among a control group. The final series had fourteen critical items distributed among thirty-two filler items. For another, shorter series of comparable items split-half reliabilities ranged from .75 to .90 for different groups. From those values Tuddenham estimated reliability to be between .80 and .93 for the series used in the present project.[16] (The critical items are illustrated in Figures 4 to 17 in the appendix.)

Subjects were presented with these tasks of visual discrimination. The equipment was operated in the manual mode: Experimenter 1 simulated the answers of four subjects on the console of each subject. On some items the simulated answers were all the modal answer established by the control group; on other items the simulated answers were divided between the mode and the next-to-mode. However, on some

---

15 Read D. Tuddenham, Studies in Conformity and Yielding X, Technical Report No. 11, Contract NR 170-159, Berkeley, University of California, 1960, 1-35 p.

16 Ibid., p. 5.

items the simulated answers were bogus, extremely deviant from the control group mode; in some cases the same answer choice was simulated four times, in other cases two adjacent answer choices were simulated twice each.

The critical items, on which subjects' responses were scored, were those where subjects were designated fifth to answer and where bogus answers were simulated. On such items a subject made her response only after perceiving the simulated answers of the four other subjects, answers which differed from the subject's own correct judgment. When a subject's response departed from the modal answer in the direction of the bogus answer, it was taken as evidence that a subject yielded her own opinion to conform to the consensus.

Every slide of the series presented nine answer choices arranged in increasing size. The answer choices differed in size by a uniform amount so that the differences between them were answer units. The measure of conforming behavior was the total number of answer units by which a subject differed from the modal answers in the direction of the bogus answers. No yielding resulted in a score of zero; complete conforming to the bogus consensus on every critical item resulted in a score of forty-three.

In a comparable laboratory procedure, but using items of different types, Crutchfield[17] found that

---

17 David Krech, Richard S. Crutchfield, and Egerton L. Ballachey, Individual in Society, New York, McGraw-Hill, 1962, p. 509-510.

considerable yielding can be produced, even when the consensus is manifestly wrong; that yielding is greater on difficult items than on easy ones; that there are large individual differences in amount of yielding; that when subjects are retested later and privately, "a major part" of the original yielding effect disappears, although there are marked individual differences in how much the effect persists.

Among situational factors affecting the Crutchfield measure of conformity, Krech et al.[18] noted the following: difficulty of task, size of group, unanimity of group consensus, size of discrepancy between individual and group, strength of coercion. These factors were held constant in the present investigation: all subjects underwent the same experimental situation. Another relevant factor, larger social context of repression or freedom, likely was very much the same for all subjects. Another factor, level of competence and acquaintance among experimental subjects, likely was of little influence since most subjects were not of close acquaintance.

---

18 Ibid., p. 512-515.

C. The Self-Report Inventory

The self-report inventory comprised the first 120 items of the MMPI, Form R,[19] excluding from tabulation fifteen items which were not suitable for adapting to error-choice format. The inventory was administered with the Form R booklet and the appropriate instructions and answer sheet. Subjects were urged to answer every item.

D. The Error-Choice Inventory

The error-choice version of personality inventory was constructed from the first 120 items of the MMPI, Form R. Excluded from tabulation were fifteen items which had an endorsement frequency too extreme (smaller than .03 or larger than .97) to allow meaningful error answer choices to be determined.[20] For each of the remaining 105 items the answer choices were derived by increasing and decreasing the true endorsement frequency by an equal amount. The size of the change depended on the size of the true endorsement frequency. Table I reports the various changes.

---

19 S.R. Hathaway and J.C. McKinley, Minnesota Multiphasic Personality Inventory, Form R, New York, Psychological Corporation, 1965, 23 p.

20 The Form R booklet numbers of items excluded are 9, 16, 23, 35, 48, 53, 54, 56, 60, 72, 85, 88, 90, 113, 114.

Table I.-

The Changes in Item Endorsement Frequencies to Form the Answer Choices for the EC Inventory.

Endorsement Frequency From	To	Change
.03 .96	.04 .97	±.03
.05 .90	.10 .95	±.04
.11 .61	.39 .89	±.07
.40	.60	±.08

The true endorsement frequencies for MMPI items were obtained from a population of 108 female college students in 1960 by Goldberg.[21] Indicative of the representativeness and generalizability of his data, endorsement frequencies from his sample at the University of Oregon correlated .97 with endorsement by eighty-five female students at Stanford University in 1959.[22] Both those samples were "highly correlated" with endorsement frequencies obtained from 125 female students at the University of Minnesota in 1962.[23] Thus the Goldberg data appeared adequate as the base from which to construct the EC answer choices.

The two error answer choices for each of the 120 MMPI statements were printed on a separate answer sheet, which also included instructions for the EC inventory. The inventory was identified as "Test of Insight into Other People," and instructions followed those used by Wilde and Fortuin.[24,25] The answer sheet was used with the Form R

---

21 Lewis R. Goldberg and Leonard G. Rorer, "Test-Retest Item Statistics for Original and Reversed MMPI Items," Oregon Research Institute Research Monograph, Vol. 3, No. 1, 1963.

22 Jerry S. Wiggins and Lewis R. Goldberg, "Interrelationships Among MMPI Item Characteristics," Educational and Psychological Measurement, Vol. 25, No. 2, 1965, p. 385.

23 Jerry S. Wiggins, Personal Correspondence with the Author, letter dated August 6, 1971.

24 G.J.S. Wilde and S. Fortuin, "Self-Report and Error-Choice: An Application of the Error-Choice Principle to the Construction of Personality Test Items," British Journal of Psychology, Vol. 60, No. 1, 1969, p. 104.

25 A copy of the instructions plus answer sheet is included in the appendix.

booklet of MMPI statements. Subjects were urged to answer every item.

For each subject an SR-EC correspondence score was derived: the number of times that a self-report item was endorsed and the counterpart error-choice item was answered with the overestimate, plus the number of times that a self-report item was answered FALSE and the counterpart error-choice item was answered with the underestimate. The range of scores possible was from zero to 105.

4. Statistical Analysis.

Before proceeding with the main statistical analysis, the independence of conformity scores and SD response set scores was tested by means of a Pearson product-moment correlation. The following formula was used for the raw scores:[26]

$$r_{xy} = \frac{N\Sigma XY - \Sigma X \Sigma Y}{\sqrt{N\Sigma X^2-(\Sigma X)^2}\sqrt{N\Sigma Y^2-(\Sigma Y)^2}}$$

For the main analysis, a three-way, fixed effects analysis of variance was used. The SR-EC correspondence scores for subjects were the dependent variable. Conformity

---

26 Lawrence-T. Dayhaw, Manuel de Statistique, Ottawa, Université d'Ottawa, 1966, p. 126.

and SD response set were classification factors with three levels each. The two orders of administration of inventories represented an experimental factor, providing a 3x3x2 complete factorial design with n equal to six per cell.[27] The analysis of variance design is illustrated in Figure 3.

5. The Specific Hypotheses.

The design, the general method of the experiment, and the statistical procedures have been described. For statistical testing, the experimental hypotheses are stated in the null form as follows:

1. There are no significant differences in the amounts of SR-EC correspondence shown by subjects who are classified as low, average and high in SD response set.

2. There are no significant differences in the amounts of SR-EC correspondence shown by subjects who are classified as low, average and high in conformity.

3. There is no significant difference in the amounts of SR-EC correspondence shown by subjects who answer the SR inventory followed by the EC version, and by subjects who answer the EC inventory followed by the SR version.

---

27 The calculation was performed by a WANG 700 computer, using a program for three-way, fixed effects analysis of variance from the Ontario Institute for Studies in Education, Toronto. The program produced the sum of squares, degrees of freedom, mean square and F value for each factor and each interaction; the sum of squares, degrees of freedom and mean square for the error term; and the total sum of squares and degrees of freedom.

	A_1			A_2			A_3		
	B_1	B_2	B_3	B_1	B_2	B_3	B_1	B_2	B_3
C_1	S_1 ⋮ S_6	S_{13} ⋮ S_{18}	S_{25} ⋮ S_{30}	S_{37} ⋮ S_{42}	S_{49} ⋮ S_{54}	S_{61} ⋮ S_{66}	S_{73} ⋮ S_{78}	S_{85} ⋮ S_{90}	S_{97} ⋮ S_{102}
C_2	S_7 ⋮ S_{12}	S_{19} ⋮ S_{24}	S_{31} ⋮ S_{36}	S_{43} ⋮ S_{48}	S_{55} ⋮ S_{60}	S_{67} ⋮ S_{72}	S_{79} ⋮ S_{84}	S_{91} ⋮ S_{96}	S_{103} ⋮ S_{108}

Legend: A SD response set
B Conformity
C Order of administration
S Subject

Figure 3.- Diagram of the Analysis of Variance Design.

4. There are no significant interactions between SD response set, conformity and order of administration of inventories.

The next chapter presents the results of the tests of significance.

CHAPTER III

PRESENTATION AND DISCUSSION OF RESULTS

In the first section this chapter restates the null hypotheses and reports the results of the statistical testing for significance. The second section presents a discussion of the meaning of the results, especially in relation to the theories advanced about error-choice inventory. Concluding the chapter are suggestions with regard to further research.

1. Results of the Tests of Significance.

The experimental design described in Chapter II required that SD response set and conformity-proneness be uncorrelated variables. The independence of the two measures used in the experiment was not conclusively demonstrated in the literature. Therefore a preliminary step was to calculate the range, mean, and standard deviation for each of the measures, and also the Pearson "r" between them. These statistics are reported in Table II. The correlation of -.03 was not significant and thus provided support for the independence of the measures. On the basis of the insignificant correlation, it was decided to use both SD response set scores and conformity scores as classification variables in the analysis of variance design.

Table II.-

Distribution and Correlation[a] of SD Response Set Scores and Conformity Scores.

Measure	N	Range From	Range To	Mean	S.D.	"r"	Significance
SD Response Set	108	14	38	28.49	5.42		
						-.03	N.S.
Conformity	108	0	42	14.53	9.73		

a Pearson "r"

The SR-EC correspondence scores of subjects were the dependent variable in the three-way, fixed effects analysis of variance which was used to test the null hypotheses of the experiment. The results are reported in Table III.[1]

However, the F values of Table III were not the final test of significance for the first three null hypotheses. A pooling procedure, as described by McNemar,[2] was used. The F value for the triple interaction of SD set, conformity, and order was .88; since it was not significant, the sum of squares for the triple interaction was pooled with the error term to derive the F ratios for the double interactions. The F values for the double interactions remained insignificant at the .01 level of probability. Therefore the sums of squares and degrees of freedom for all interactions were pooled with the original sum of squares and degrees of freedom, respectively, for the error term. The pooled error term was used to check the variance of the main effects, as reported in Table IV.

The first hypothesis of the experiment in its null form states that there are no significant differences in

---

1 The means for the various combinations of levels of factors are listed in the appendix.

2 Quinn McNemar, Psychological Statistics, New York, Wiley, 1969, p. 382-383.

Table III.-

Three-Way, Fixed Effects Analysis of Variance of SR-EC Correspondence Scores for Different Levels of SD Response Set and Conformity and Different Orders of Administration.

Source	SS	df	MS	F	Sig.
SD Response Set	29.56	2	14.78	.19	N.S.
Conformity	28.50	2	14.25	.18	N.S.
Order	277.12	1	277.12	3.50	N.S.[a]
SD x Conformity	76.78	4	19.19	.24	N.S.
SD x Order	193.85	2	96.93	1.22	N.S.
Conformity x Order	19.68	2	9.84	.12	N.S.
SD x Conformity x Order	279.26	4	69.82	.88	N.S.
Error	7136.17	90	79.29		
Total	8040.92	107			

a For .05 level of significance, $F(1,100) \geq 3.94$.

Table IV.-

Analysis of Main Effects Variance of SR-EC Correspondence Scores, Using a Pooled Error Term.

Source	SS	df	MS	F	Sig.
SD Response Set	29.56	2	14.78	.20	N.S.
Conformity	28.50	2	14.25	.18	N.S.
Order	277.12	1	277.12	3.67	N.S.[a]
Pooled Error	7705.74	102	75.55		

a For .05 level of significance, $F(1,100) \geq 3.94$.

the amounts of SR-EC correspondence shown by subjects who are classified as low, average and high in SD response set.

From the analysis of variance the F value for the levels of SD response set was .20, which was not significant at the .01 level of probability. Therefore the first null hypothesis was not rejected.

The second hypothesis in its null form states that there are no significant differences in the amounts of SR-EC correspondence shown by subjects who are classified as low, average and high in conformity.

The F value for the levels of conformity was .18, which was not significant at the .01 level. Therefore the second hypothesis was not rejected.

The third hypothesis of the experiment in its null form states that there is no significant difference in the amounts of SR-EC correspondence shown by subjects who answer the SR inventory followed by the EC version, and by subjects who answer the EC inventory followed by the SR version.

The F value for the two orders of administration of the inventories was 3.67, which was not significant at the .01 level of probability. Consequently, the third hypothesis was not rejected.

The fourth hypothesis of the experiment in its null form states that there are no significant interactions

between SD response set, conformity, and order of administration of inventories.

It was noted above that the F value for the interaction of SD set x conformity x order was .88, which was not significant. Table III also reports the F value for SD set x conformity interaction as .24; that for SD set x order interaction as 1.22; and that for conformity x order interaction as .12. None of these values was significant at the .01 level. Therefore, the fourth hypothesis was not rejected.

2. Discussion of Results.

Individual differences in the tendency to describe oneself in socially desirable terms were not a significant source of variance in SR-EC correspondence. That finding provides some support for the proposal by Wilde and his colleagues[3] that the error-choice format of personality inventory is able to eliminate the influence of SD response set. It is also support for Edwards'[4] definition of SD response set, which restricted the influence of social desirability to self-descriptions.

---

3 G.J.S. Wilde and S. Fortuin, "Self-Report and Error-Choice: An Application of the Error-Choice Principle to the Construction of Personality Test Items," British Journal of Psychology, Vol. 60, No. 1, 1969, p. 102.

4 Allen L. Edwards, "The Social Desirability Variable: A Broad Statement," in Irwin A. Berg (ed.), Response Set in Personality Assessment, Chicago, Aldine, 1967, p. 46-47.

The failure to find significance for the SD response set variable does not support the interpretation of factor analyses offered by Kusyszyn and Jackson[5] and by Jackson and Messick.[6] These authors interpreted a social desirability factor in subjects' judgments of endorsement frequencies to be the same social desirability factor that exists in subjects' self-descriptions. It seems appropriate to recall Rorer's[7] caution that factor analytic results might be consistent with a response style interpretation but that the results would not be proof of the existence of response styles.

With regard to the decision not to reject the null hypothesis about SD response set, however, it is worth noting that the power of the F test[8] was less than .20 for alpha equal to .01. In other words, the probability was less than one in five of rejecting the null hypothesis if the

---

5 Igor Kusyszyn and Douglas N. Jackson, "A Multimethod Factor Analytic Appraisal of Endorsement and Judgment Methods in Personality Assessment," Educational and Psychological Measurement, Vol. 28, No. 4, 1968, p. 1047-1061.

6 Douglas N. Jackson and Samuel Messick, "A Distinction Between Judgments of Frequency and of Desirability as Determinants of Response," Educational and Psychological Measurement, Vol. 29, No. 2, 1969, p. 273-293.

7 Leonard G. Rorer, "The Great Response-Style Myth," Psychological Bulletin, Vol. 63, No. 3, 1965, p. 129-156.

8 Gene V. Glass and Julian C. Stanley, Statistical Methods in Education and Psychology, Englewood Cliffs, N.J., Prentice-Hall, 1970, p. 376-377.

actual differences in means in the population were of the size manifested by the present sample. The power of the F tests in the experiment will receive additional comment after all the null hypotheses have been reviewed.

The second hypothesis of the present investigation was developed from a rationale proposed by Wilde and Fortuin.[9] It theorized that subjects make their self-report responses conform to their estimate of the population norm. Error-choice responses would reflect the same estimate of the population, thus accounting for the SR-EC correspondence. The absence of significant differences in SR-EC correspondence for the three levels of conformity-proneness does not support the hypothesized rationale.

The null hypotheses concerning socially desirable responding and conformity were not rejected by the tests of significance. However, additional research might investigate the constructs of SD response set and conformity with measuring instruments different from those used in the present thesis. Edwards' SD scale was used because Edwards was the major author in the area of SD response set. Wilde and his co-authors proposed the rationale based on conformity, but they did not define conforming behavior. The present investigator chose the Crutchfield-Tuddenham

---

9 Wilde and Fortuin, "Self-Report and Error-Choice: An Application of the Error-Choice Principle to the Construction of Personality Test Items," p. 103.

measure although the literature revealed disagreement about definition and about the generalizability of measures.

On one side of the issue Asch[10] warned against considering conformity "a psychologically homogeneous phenomenon," maintaining that conforming behavior could result from different processes. Steiner and Vannoy[11] suggested that the conformity exhibited in the laboratory was not homogeneous. They discriminated two groups among laboratory subjects, those who reasserted their conforming responses several days later in a private situation, and those who did not. McDavid and Sistrunk[12] obtained results in the Crutchfield situation which yielded a low correlation between conformity on easy tasks and conformity on extremely difficult tasks, suggesting the two were different processes. Blake and Mouton[13] concluded that it is difficult to generalize laboratory performance to more general and lifelike

---

10 Solomon E. Asch, "Issues in the Study of Social Influences on Judgment," in Irwin A. Berg and Bernard M. Bass (eds.), Conformity and Deviation, New York, Harper, 1961, p. 154.

11 Ivan D. Steiner and Joseph S. Vannoy, "Personality Correlates of Two Types of Conformity Behavior," Journal of Personality and Social Psychology, Vol. 4, No. 3, 1966, p. 307-315.

12 John W. McDavid and Frank Sistrunk, "Personality Correlates of Two Kinds of Conforming Behavior," Journal of Personality, Vol. 32, 1964, p. 420-435.

13 Robert R. Blake and Jane Srygley Mouton, "Conformity, Resistance and Conversion," in Irwin A. Berg and Bernard M. Bass (eds.), Conformity and Deviation, New York, Harper, 1961, p. 5.

situations because the laboratory measures of conformity are artificial and possess little intrinsic interest.

In contrast to those views, other authors argued in favor of a consistent trait of conformity. Chipman[14] reported that subjects showed a consistent readiness to exhibit conformity and did not show a differential responsiveness to experimental variables of social pressure and judgment difficulty. Back and Davis[15] found modest but significant ($p < .05$) correlations between three measures of conformity: the Crutchfield task, agreement with peers' standards, and agreement with authority's regulations. With reference to the generalizability of a laboratory measure, Krech et al.[16] theorized that salient aspects of the laboratory situation are also true of life situations. In life situations the individual often does not have the opportunity to engage in discussion of the issue with others; he often cannot search out relevant facts, perhaps because the issue is too complex; he often cannot withhold judgment if he is

---

14 Abram Chipman, "Conformity as a Differential Function of Social Pressure and Judgment Difficulty," Journal of Social Psychology, Vol. 70, No. 2, 1966, p. 307.

15 Kurt W. Back and Keith E. Davis, "Some Personal and Situational Factors Relevant to the Consistency and Prediction of Conforming Behavior," Sociometry, Vol. 28, No. 3, 1965, p. 227-240.

16 David Krech, Richard S. Crutchfield and Egerton L. Ballachey, Individual in Society, New York, McGraw-Hill, 1962, p. 512.

unsure. Hicks et al.[17] modified laboratory conditions, making them more analogous to those often found in daily life: when the bogus group was removed from the subject's presence, Asch-type group influence still operated.

Notwithstanding the differing opinions, the Crutchfield-Tuddenham procedure does give a sample of behavior which it seems reasonable to label "conforming to group opinion." Self-report measures of conformity would allow possible inaccuracy due to response sets, and they are, therefore, not preferable to a direct measure of behavior. Moreover, the measure of conformity used in the experiment was chosen to be comparable to the conformity hypothesized to occur when individuals answer self-report inventories. In the latter case there is no particular, direct pressure to conform. Similarly, in the Crutchfield-Tuddenham procedure there were no special, direct pressures to conform. Subjects were participating in a research project conducted by an experimenter not acquainted with them; their answers were anonymous to other subjects; they were told that the experimenter was not concerned with individual results. Thus in both the Crutchfield procedure and in responses to inventories the conforming behavior fulfills

---

17 Jack M. Hicks, Richard A. Monty and Thomas I. Myers, "Group Consensus and Judgmental Accuracy: Extension of the Asch Effect," Psychonomic Science, Vol. 5, No. 4, 1966, p. 159-160.

the definition offered by Krech et al.:[18] the yielding of one's opinions or values because of differing opinions or values held by the society or the group. The Crutchfield-Tuddenham measure, in summary, seemed appropriate for the investigation of Wilde and Fortuin's rationale about self-report inventories.

The experimental results showed no significant contribution to SR-EC correspondence variance by levels of conformity-proneness. However, Wilde and de Wit[19] allowed that the process of conforming in SR inventory might account for SR-EC correspondence in some subjects, while another process might operate to account for correspondence in other subjects. The alternative process was that of conceiving others in one's own image, making others like oneself. If one process operated in subjects high in conformity-proneness and the other process operated in subjects low in conformity-proneness, there might be no significant difference in SR-EC correspondence scores for the two groups.

In order to investigate the possibility of different processes occurring in different subjects, a post-experiment

---

18 Krech et al., Individual in Society, p. 506.

19 G.J.S. Wilde and O.E. de Wit, "Self-Report and Error-Choice: Inter-Individual Differences in the Operation of the Error-Choice Principle and Their Validity in Personality Questionnaire Tests," British Journal of Psychology, Vol. 61, No. 2, 1970, p. 227.

questionnaire was prepared.[20] It provided brief descriptions of the two processes that might operate during responding to self-report and error-choice inventories. Subjects were asked whether one process was occurring more than the other when they completed the inventories. The questionnaire was distributed to the third of the sample high in conformity-proneness and to the low third. It was keyed for conformity level, and subjects were told they could return it anonymously.

The two groups showed little difference in the frequency with which they reported the two processes occurring. One subject in each group acknowledged that she generally made her self-descriptions conform to the population norm. Thirty conformists and thirty-two independents reported they generally made others like themselves; five conformists and three independents reported that neither of the two processes occurred during their responding. A X^2 was calculated for these frequencies; the value of .58 was not significant, for two degrees of freedom. The finding does not support the theory that a conformity process might operate in subjects high in conformity-proneness and a different process in others. The influence of level of conformity on SR-EC correspondence remains undemonstrated.

---

20 A copy of the post-experiment questionnaire is included in the appendix.

However, a remark is in order about the power of the F test for the null hypothesis concerning conformity-proneness. For alpha equal to .01 the power of the test was less than .15; that is, there was less than one chance in five of rejecting the null hypothesis if the actual differences in means in the population were of the size manifested by the present sample.

The writer offers a final comment on the hypothesized conformity motivation in SR responding. If subjects did make their self-descriptions conform to their estimate of the population norm, it would not follow that their true opinions and their behaviors conformed to the norm. To describe oneself like the norm could be considered a response style or a form of impression management, which would reduce the accuracy of SR trait scores as measures of a subject's personality. If EC inventory responding did correspond to conformity-motivated SR responding, the error-choice responses would reflect distorted self-descriptions; they would not have increased validity as measures of a subject's personality. If the process of conforming in SR inventory accounted for SR-EC correspondence, the error-choice format of inventory would be inaccurate as a method of personality assessment.

The discussion now turns to order of administration of inventories as a factor in the amount of SR-EC

correspondence. The data showed a tendency for subjects who complete the EC inventory first to show a larger amount of SR-EC correspondence than subjects who complete the SR inventory first. However, the F test did not result in a significant value and the null hypothesis was not rejected. For alpha equal to .01, the F test had a power of less than .45. Conversely, the probability of a Type II error was greater than .55: the probability was greater than one in two that the null hypothesis would be accepted when the observed differences in means represented the true differences in the population.

The largest F values from the analysis of variance were those for order of administration and for SD set x order interaction. Therefore it was decided to perform an analysis of variance for the simple effects[21] of the orders of administration at each level of SD response set. The tests on simple effects are reported in Table V.

The only F value that approached significance was that for the simple effects of the orders of administration for subjects low in SD response set. The probability is less than .05 that an F value of 4.63 would occur by chance. The results of the analysis of variance for simple effects

---

21 Lawrence-T. Dayhaw, Manuel de Statistique, Ottawa, Université d'Ottawa, 1966, p. 464-465.

Table V.-

Analysis of Variance for Simple Effects of Orders of Administration for Each Level of SD Response Set.

Source	SS	df	MS	F	Sig.
Order for low SD	367.36	1	367.36	4.63	.05[a]
Order for average SD	.25	1	.25	.003	N.S.
Order for high SD	103.36	1	103.36	1.30	N.S.
Error	7136.17	90	79.29		

a For .01 level of significance, $F(1,80) \geq 6.96$.

suggest that if the F test had had increased power the over-all main effects of order of administration might have proven significant. In any case, the order of administration of inventories should be controlled in any further investigation of SR-EC correspondence.

In subsequent research it would be desirable to increase the power of the F tests. In the present experiment, the test of the null hypothesis concerning order of administration had a power less than .45, and the power of the other tests for main effects variance was smaller still. Increasing the size of the sample would serve to increase the power of the tests, as noted by Kirk.[22]

The results of the tests of significance can tentatively be extended in meaning. Since SD response set did not prove a significant factor in SR-EC correspondence, it suggests that SD set is not equally influential in EC inventory as in SR inventory. In addition, the large majority of subjects acknowledged that they sometimes estimated endorsement frequencies by conceiving others in their own image. These findings combine to give tentative support to the basic theory that error-choice inventory can reveal the personality traits of the subject without the

---

22 Roger E. Kirk, Experimental Design: Procedures for the Behavioral Sciences, Belmont, Cal., Wadsworth, 1968, p. 10.

contaminating variance from SD response set. Further research might profitably be directed toward investigating the validity of trait scores derived from EC inventory.

The present experiment has not tested Wilde and Fortuin's[23] first rationale for SR-EC correspondence. The rationale provided that a subject makes others like himself in order to resolve dissonance between SD motivation and either honesty or modesty motivation. The acknowledgment by many subjects that they conceived other people in their own image when making EC responses is consistent with the rationale. However, it is not evidence for the hypothesized process of conflicting motivations--cognitive dissonance--adjustment of cognitions. It may be that subjects conceive others in their own image without any reference to self-report responses, SD motivation, etc.

The present report does confirm that the SR-EC correspondence hypothesized by Wilde and his colleagues is more than a chance phenomenon. The correspondence occurred in .71 of the items; it could be expected in .50 of the items by chance. A t test for the difference between those proportions yielded a value of 33.33, significant at the .001 level of probability. The error-choice format of

---

23 Wilde and Fortuin, "Self-Report and Error-Choice: An Application of the Error-Choice Principle to the Construction of Personality Test Items," p. 103.

inventory appears worthy of further theory and investigation.

In anticipation of subsequent research, the writer raises a final practical question. Is the error-choice version of personality inventory as widely applicable as self-report inventory? With its dependence on percentages, the EC version might require a higher minimum level of intelligence or education for respondents.

SUMMARY AND CONCLUSIONS

The error-choice format of personality inventory had been proposed as an indirect method of obtaining trait scores free from distortion by social desirability response set. Rationales had been advanced to account for correspondence between responses to a self-report personality inventory and responses to an error-choice version. The present experiment sought to investigate the influence of three factors on the amount of SR-EC correspondence.

The first hypothesis stated that there are no significant differences in the amounts of SR-EC correspondence shown by subjects who are classified as low, average and high in SD response set. The second hypothesis stated that there are no significant differences in the amounts of SR-EC correspondence shown by subjects who are classified as low, average and high in conformity. The third hypothesis stated that there is no significant difference in the amounts of SR-EC correspondence shown by subjects who answer the SR inventory followed by the EC version, and by subjects who answer the EC inventory followed by the SR version. The fourth hypothesis stated that there are no significant interactions between SD response set, conformity, and order of administration of inventories.

A three-way fixed effects analysis of variance solution was used as the test of significance for the hypotheses.

The levels of SD response set and conformity did not produce significant results; therefore the first two null hypotheses were not rejected. The lack of significant findings for the SD factor was considered tentative support for the theory that error-choice inventory eliminates the influence of SD response set. The absence of significant differences for conformity levels did not support the rationale that a conformity motivation in self-report inventory accounts for SR-EC correspondence.

The two orders of administration of the inventories did not yield significant differences. However, an F test for the simple effects of orders for subjects low in SD set approached significance. Thus the third hypothesis was not rejected, but the order variable was considered to warrant control in subsequent research. It was noted that the F test had a low power, which meant a low probability of finding significant results.

None of the interactions yielded significant results, and the fourth hypothesis was not rejected.

It was suggested that further research could profitably be directed toward the processes underlying SR-EC correspondence, and toward the validation of trait scores derived from error-choice inventory.

BIBLIOGRAPHY

Crutchfield, Richard S., "Conformity and Character," American Psychologist, Vol. 10, No. 5, 1955, p. 191-198.
The author reports results from a laboratory procedure designed to emulate Asch's situation of group influence. The Crutchfield procedure, as modified by Tuddenham et al., was the measure of conformity in the present experiment.

Dahlstrom, W. Grant, "Commentary: The Roles of Social Desirability and Acquiescence in Responses to the MMPI," in Samuel Messick and John Ross (eds.), Measurement in Personality and Cognition, New York, Wiley, 1962, p. 157-168.
The author questions the conclusions drawn from statistical results; he argues that trait scores may be valid measures despite confounding by response sets.

Edwards, Allen L., The Social Desirability Variable in Personality Assessment and Research, New York, Dryden, 1957, viii-108 p.
The author develops his position that the social desirability continuum is the most important dimension of personality statements. He describes the construction of the scale used in the present thesis to measure SD response set.

--------, "The Social Desirability Variable: A Broad Statement," in Irwin A. Berg (ed.), Response Set in Personality Assessment, Chicago, Aldine, 1967, p. 32-47.
Edwards is a major theorist and researcher in the area of response sets. Here he defines SD set as a personality trait, and he illustrates the relationship between the social desirability scale value of items and their probability of endorsement.

--------, The Measurement of Personality Traits by Scales and Inventories, New York, Holt, Rinehart and Winston, 1970, xiii-306 p.
Edwards cites many studies and presents clear restatements of his arguments that inventory trait scores are confounded by SD response set. Other chapters provide useful background in the theory and construction of inventories.

Edwards, Allen L., Carol J. Diers, and Jerald N. Walker, "Response Sets and Factor Loadings on Sixty-One Personality Scales," Journal of Applied Psychology, Vol. 46, No. 3, 1962, p. 220-225.
The study is characteristic of the factor analytic approach to the meaning of personality scale scores. Results

are interpreted to mean that much of the variance is attributable to response sets of social desirability and acquiescence.

Fiske, Donald W., and Pamela H. Pearson, "Theory and Techniques of Personality Measurement," Annual Review of Psychology, Vol. 21, 1970, p. 49-86.

This recent review of the area of personality measurement acknowledges that the existence and the influence of response sets remain in dispute. The reviewers accept their existence but minimize their importance.

Goldberg, Lewis R., and Leonard G. Rorer, "Test-Retest Item Statistics for Original and Reversed MMPI Items," Oregon Research Institute Research Monograph, Vol. 3, No. 1, 1963.

This monograph is a compilation of such MMPI item data as endorsement frequencies and social desirability scale values from various samples. The endorsement frequencies were used to construct the answer choices for the EC inventory.

Hammond, Kenneth R., "Measuring Attitudes by Error-Choice: An Indirect Method," Journal of Abnormal and Social Psychology, Vol. 43, No. 1, 1948, p. 38-49.

The author describes the innovation of the error-choice technique as a method of attitude survey. Results of his experiments suggest that the technique eliminates response sets.

Jackson, Douglas N., and Samuel Messick, "Response Styles and the Assessment of Psychopathology," in Samuel Messick and John Ross (eds.), Measurement in Personality and Cognition, New York, Wiley, 1962, p. 129-155.

The authors present an early formulation of their concept of response styles, which emphasizes wide generalizability. On the basis of a factor analysis they conclude that response styles confound the interpretation of MMPI scale scores.

Krech, David, Richard S. Crutchfield, and Egerton L. Ballachey, Individual in Society, New York, McGraw-Hill, 1962, xii-564 p.

The last chapter summarizes Crutchfield's studies of conformity-independence. The personality dimension is defined, and characteristics of subjects at each pole are reported. Factors relevant to Crutchfield's measuring procedure are noted.

Kubany, Albert J., "A Validation Study of the Error-Choice Technique Using Attitudes on National Health Insurance," Educational and Psychological Measurement, Vol. 13, No. 2, 1953, p. 157-163.

A study of the error-choice technique concludes that it is a subtle method of measuring attitudes.

Kusyszyn, Igor, and Douglas N. Jackson, "A Multi-method Factor Analytic Appraisal of Endorsement and Judgment Methods in Personality Assessment," Educational and Psychological Measurement, Vol. 28, No. 4, 1968, p. 1047-1061.

The authors interpret the results of a factor analysis to mean that social desirability response set operates in judgments of endorsement frequency. Such an interpretation contradicts the theory that error-choice inventory eliminates SD response set.

McGee, Richard K., "Response Set in Relation to Personality: An Orientation," in Irwin A. Berg (ed.), Response Set in Personality Assessment, Chicago, Aldine, 1967, p. 1-31.

In a succinct but comprehensive review of the concept of set in psychological theory, McGee cites evidence from both sides of the controversy about the importance of response sets in personality inventories. It is a valuable overview.

Rorer, Leonard G., "The Great Response-Style Myth," Psychological Bulletin, Vol. 63, No. 3, 1965, p. 129-156.

On logical grounds, Rorer criticizes studies that conclude trait score variance is attributable to response styles. In addition, he cites convincing studies which refute the interpretations of acquiescence and social desirability styles.

Rosen, Ephraim, "Self-Appraisal, Personal Desirability, and Perceived Social Desirability of Personality Traits," Journal of Abnormal and Social Psychology, Vol. 52, No. 2, 1956, p. 151-158.

Using data for the sample as a whole, Rosen reports that self-report responses to the MMPI are highly correlated with responses under instructions to endorse desirable content.

Tuddenham, Read D., "Correlates of Yielding to a Distorted Group Norm," Journal of Personality, Vol. 27, No. 2, 1959, p. 272-284.

The author reports very few significant correlations between his laboratory measure of conformity and various

personality trait scales, especially for females. The results suggest that conforming behavior is not well measured by inventory scales.

--------, Studies in Conformity and Yielding X, Technical Report No. 11, Contract NR 170-159, Berkeley. University of California, 1960, 35 p.

The series of visual discrimination items used in the present experiment is described. The series was designed to be suitable for different age groups and different cultures. Normative data for control and experimental subjects are included.

Tuddenham, Read D., Philip Macbride and Victor Zahn, Studies in Conformity and Yielding I, Technical Report No. 1, Contract NR 170-159, Berkeley, University of California, 1956, iii-72 p.

In the first report of their extended research project, the authors give a detailed description of the laboratory apparatus and the procedure for the conformity measure used in the present thesis. They note the criteria for selection of their original pool of items.

Weschler, Irving R., "An Investigation of Attitudes Toward Labor and Management by Means of the Error-Choice Method: I," Journal of Social Psychology, Vol. 32, 1950, p. 51-62.

The author's studies of the error-choice technique tend to confirm Hammond's original hypothesis that the direction of the error would be constant and related to attitudes.

Wilde, G.J.S., and O.E. de Wit, "Self-Report and Error-Choice: Inter-Individual Differences in the Operation of the Error-Choice Principle and Their Validity in Personality Questionnaire Tests," British Journal of Psychology, Vol. 61, No. 2, 1970, p. 219-228.

This paper provides additional statistical description of the experiment conducted by Wilde and Fortuin. Trait scores derived from error-choice inventory show moderate validity in relation to a gross criterion measure.

Wilde, G.J.S., and S. Fortuin, "Self-Report and Error-Choice: An Application of the Error-Choice Principle to the Construction of Personality Test Items," British Journal of Psychology, Vol. 60, No. 1, 1969, p. 101-108.

This article and the preceding one were the impetus for the present thesis. The authors adapt the error-choice technique for use with personality inventory and they offer theories to account for its operation. They suggest the technique eliminates SD response set.

APPENDIX 1

THE INVENTORY MEASURING SD RESPONSE SET

INTRODUCTORY INVENTORY

Read the following statements and decide whether each is true as applied to you or false as applied to you. Mark your answers with an X in either the true (T) or the false (F) box on the answer sheet.

Give your own opinion of yourself. Try to give an answer for every statement if possible.

1. My hands and feet are usually warm enough.
2. I am very seldom troubled by constipation.
3. I find it hard to keep my mind on a task or job.
4. Most any time I would rather sit and daydream than to do anything else.
5. My family does not like the work I have chosen (or the work I intend to choose for my life work).
6. My sleep is fitful and disturbed.
7. I am liked by most people who know me.
8. I am happy most of the time.
9. Criticism or scolding hurts me terribly.
10. It makes me impatient to have people ask my advice or otherwise interrupt me when I am working on something important.
11. I have had periods in which I carried on activities without knowing later what I had been doing.

12. I cry easily.
13. I do not tire quickly.
14. I am not afraid to handle money.
15. It makes me uncomfortable to put on a stunt at a party even when others are doing the same sort of things.
16. I frequently notice my hand shakes when I try to do something.
17. It does not bother me particularly to see animals suffer.
18. I dream frequently about things that are best kept to myself.
19. My parents and family find more fault with me than they should.
20. I have reason for feeling jealous of one or more members of my family.
21. No one cares much what happens to you.
22. I usually expect to succeed in things I do.
23. I sweat very easily even on cool days.
24. When in a group of people I have trouble thinking of the right things to talk about.
25. I can easily make other people afraid of me, and sometimes do for the fun of it.
26. I am never happier than when alone.
27. Life is a strain for me much of the time.
28. I am easily embarrassed.

29. I cannot keep my mind on one thing.

30. I feel anxiety about something or someone almost all the time.

31. I have been afraid of things or people that I knew could not hurt me.

32. I am not unusually self-conscious.

33. People often disappoint me.

34. I feel hungry almost all of the time.

35. I worry quite a bit over possible misfortunes.

36. It makes me nervous to have to wait.

37. I blush no more often than others.

38. I shrink from facing a crisis or difficulty.

39. I sometimes feel that I am about to go to pieces.

ANSWER SHEET FOR INTRODUCTORY INVENTORY

NAME _____

PAGE 1

	T	F

PAGE 2

	T	F
12		
13		
14		
15		
16		
17		
18		
19		
20		
21		
22		
23		
24		
25		
26		
27		
28		

PAGE 3

	T	F
29		
30		
31		
32		
33		
34		
35		
36		
37		
38		
39		

APPENDIX 2

THE KEYING AND THE CORRESPONDING MMPI BOOKLET NUMBER OF THE ITEMS OF THE "INTRODUCTORY INVENTORY"

APPENDIX 2

Table VI.-

The Keying and the Corresponding MMPI Booklet Number of the Items of the "Introductory Inventory."

Item	Keyed Direction	Booklet Number	Item	Keyed Direction	Booklet Number
1	True	7	21	False	252
2	True	18	22	True	257
3	False	32	23	False	263
4	"	40	24	"	267
5	"	42	25	"	269
6	"	43	26	"	286
7	True	54	27	"	301
8	True	107	28	"	321
9	False	138	29	"	335
10	"	148	30	"	337
11	"	156	31	"	352
12	"	158	32	True	371
13	True	163	33	False	383
14	True	169	34	"	424
15	False	171	35	"	431
16	"	186	36	"	439
17	"	218	37	True	528
18	"	241	38	False	549
19	"	245	39	False	555
20	"	247			

APPENDIX 3

INSTRUCTIONS TO SUBJECTS IN THE PROCEDURE FOR MEASURING CONFORMITY

APPENDIX 3

INSTRUCTIONS TO SUBJECTS IN THE PROCEDURE FOR MEASURING CONFORMITY[1]

Thank you for coming today to help me. First, have a look at the answering panel in front of you. There is a row of switches at the bottom numbered 1 to 9; red lights on the left side, numbered 1 to 5; and rows of green lights opposite them. (Room light off.)

Now let me describe what we are going to do. I'm going to show you some slides on the screen. Each slide is a problem. This is an example. (Practice slide on.) On this one the question is, "Which of the numbered lines is slanted the same as the line marked C?" You can see that there are nine lines and that each line has a number. Now look at the panel in front of you. At the bottom, there are nine switches numbered 1 to 9. Later on, you will use these switches to give your answers. Each slide has nine alternative answers. So, you will indicate your answer by turning on the switch whose number corresponds to the alternative you choose. (Slide off.) I want you to be as <u>accurate</u> as you can.

Let's try it out to see how it works. You must answer <u>in turn</u> so that the apparatus can record accurately. (Automatic mode, power on, answering order 2-4-1-5-3.) The column

---

1 Adapted from Read D. Tuddenham, <u>et al.</u>, <u>Studies in Conformity and Yielding I</u>, Technical Report No. 1, Contract NR 170-159, Berkeley, University of California, 1960, p. 64-67.

of red lights is numbered 1 to 5 from top to bottom. When the apparatus is turned on, one of the lights will be red in your booth. The red light that is on in your booth indicates your order in answering. When the red light is on in row one, you answer first; when it is on in row two, you answer second, and so on. You will answer in a different order on each trial.

Miss _____, your red light is on in row one, so you are to answer first this time. Just for practice, will you pull a switch toward you--any switch will do. Tell me which switch you pulled. ..... Thank you. Now look at your panels. Green light number ____ should have come on in the first row on all the panels. This green light signals the person second in order to give her answer next. This time it is Miss _____ whose red light is on in row two. Will you please pull some other switch and tell me which one you chose. ..... Thank you. This time green light number _____ came on in row two. This signals Miss _____ who is third in turn, to pull a switch. Will you choose still another number and state your choice aloud. ..... Thank you. Green light number _____ came on in the third row. This is the signal for Miss _____, whose red light is on in the fourth row, to pull one of the switches. Also state your number aloud..... Thank you. Green light number _____ came on in the fourth row. It is now your turn,

Miss _____ . ..... This time green light number _____ came on in the fifth row.

Your switches should still be on. The signal to release them is when the panel lights go out. Please move your switches gently by holding them between your thumb and finger. Please don't snap them on or off. This is necessary for proper recording and it also prevents damage to the apparatus. (Power off.) Please turn off your switches.

Now let's briefly go over what has happened. The red light tells you which will be your turn. The green lights tell you when to take that turn. So, you must first watch the red light to see which is your turn, and, second, watch the green lights to see when it is time to take your turn. For example, if red light number three is on on your panel, you are third in turn and you are to wait until the green lights come on in rows one and two before you turn on your switch to answer.

Now we will have some practice slides for you to get used to the apparatus. Go right ahead and give your answer when your turn comes up. (Three practice slides.) .....

Are there any questions about how to run the apparatus? We want to finish as quickly as possible, so I will ask that there be no questions or talking from now until we finish. I think you can give your answers without too much trouble and we can move right along.

Please remember to move your switches carefully--do not snap them. Turn them off when the lights go off. Make your answers as accurate as you can. (Change to manual mode.)

APPENDIX 4

THE CRITICAL ITEMS IN THE PROCEDURE FOR MEASURING CONFORMITY

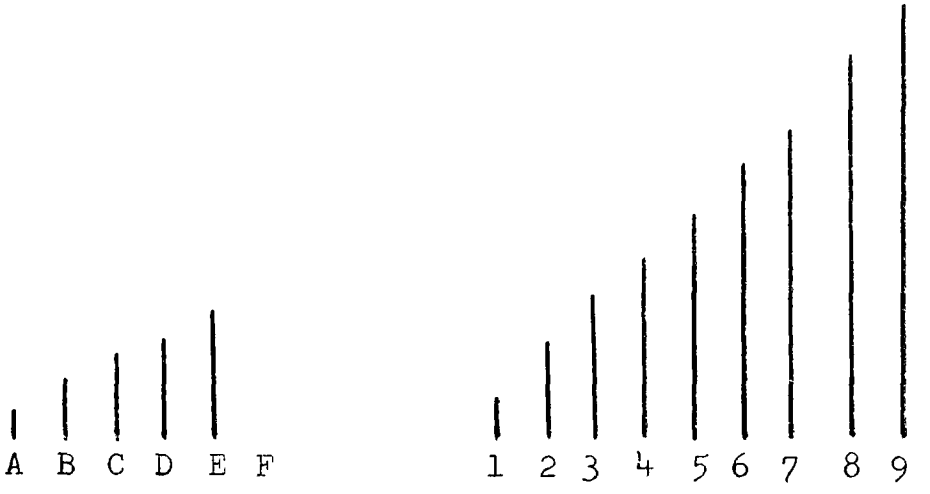


Figure 4.- Diagram of the First Critical Item[a] in the Procedure for Measuring Conformity-Proneness.

Read D. Tuddenham, Studies in Conformity and Yielding X, Technical Report No. 11, Contract NR 170-159, Berkeley, University of California, 1960, p. 6.

a The accompanying verbal statement of the task was "Look at the lines with letters and see how they go. One of the numbered lines should come next. Which of the numbered lines belongs at F?" The answers simulated were 7,6,7,6.

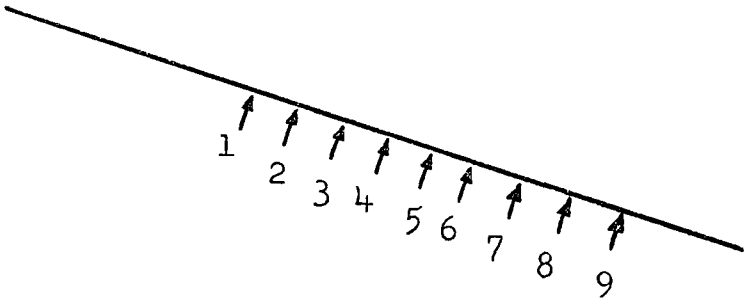


Figure 5.- Diagram of the Second Critical Item[a] in the Procedure for Measuring Conformity-Proneness.

Read D. Tuddenham, Studies in Conformity and Yielding X, Technical Report No. 11, Contract NR 170-159, Berkeley, University of California, 1960, p. 6.

a The accompanying verbal statement of the task was "Which of the numbered arrows points to the exact middle of the long line?" The answers simulated were 7,7,7,7.

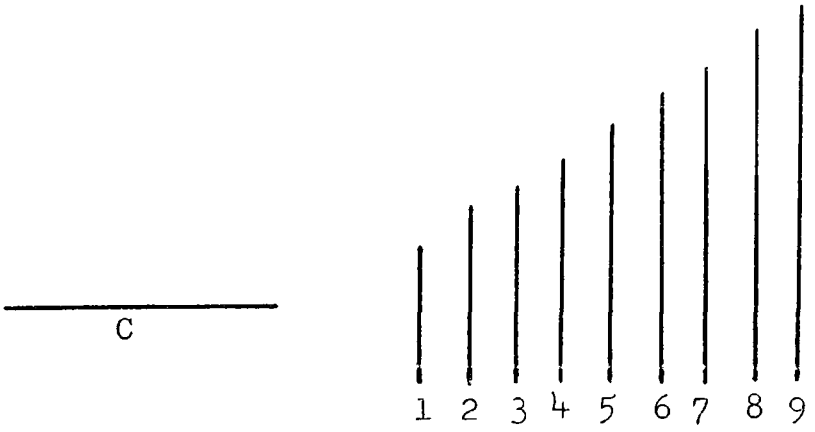


Figure 6.- Diagram of the Third Critical Item[a] in the Procedure for Measuring Conformity-Proneness.

Read D. Tuddenham, Studies in Conformity and Yielding X, Technical Report No. 11, Contract NR 170-159, Berkeley, University of California, 1960, p. 6.

a The accompanying verbal statement of the task was "One of the numbered lines is the same length as the line marked C. Which line is it?" The answers simulated were 6,7,6,7.

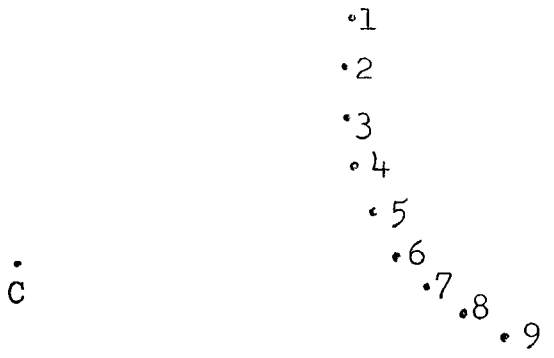


Figure 7.- Diagram of the Fourth Critical Item[a] in the Procedure for Measuring Conformity-Proneness.

Read D. Tuddenham, Studies in Conformity and Yielding X, Technical Report No. 11, Contract NR 170-159, Berkeley, University of California, 1960, p. 6.

a The accompanying verbal statement of the task was "Which of the numbered dots is closest to the dot marked C?" The answers simulated were 7,6,6,7.

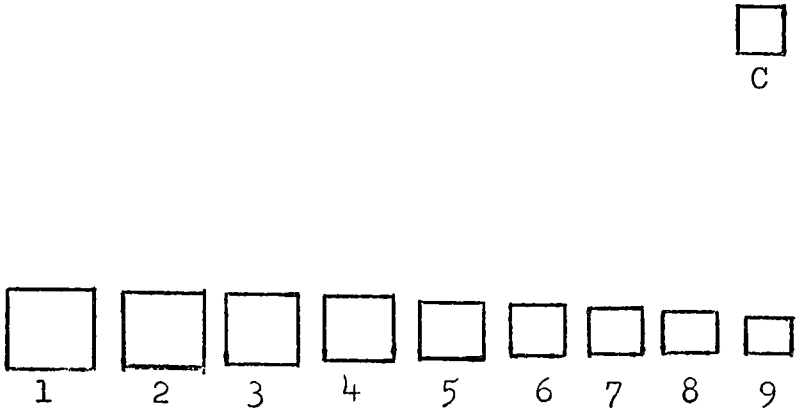


Figure 8.- Diagram of the Fifth Critical Item[a] in the Procedure for Measuring Conformity-Proneness.

Read D. Tuddenham, Studies in Conformity and Yielding X, Technical Report No. 11, Contract NR 170-159, Berkeley, University of California, 1960, p. 6.

a The accompanying verbal statement of the task was "Here are some squares of different sizes. Which numbered square is the same size as the square marked C?" The answers simulated were 4,3,3,4.

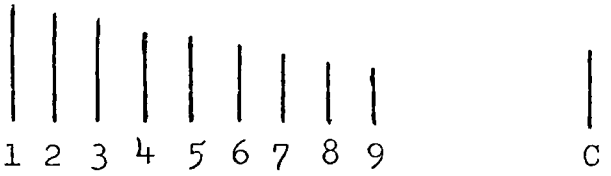


Figure 9.- Diagram of the Sixth Critical Item[a] in the Procedure for Measuring Conformity-Proneness.

Read D. Tuddenham, Studies in Conformity and Yielding X, Technical Report No. 11, Contract NR 170-159, Berkeley, University of California, 1960, p. 6.

a The accompanying verbal statement of the task was "One of the numbered lines is the same length as the line marked C. Which line is it?" The answers simulated were 3,2,2,3.

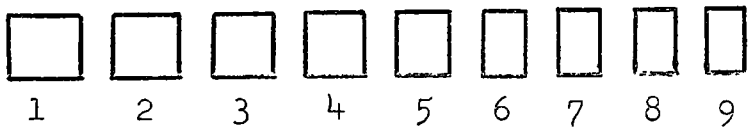


Figure 10.- Diagram of the Seventh Critical Item[a] in the Procedure for Measuring Conformity-Proneness.

Read D. Tuddenham, Studies in Conformity and Yielding X, Technical Report No. 11, Contract NR 170-159, Berkeley, University of California, 1960, p. 6.

a The accompanying verbal statement of the task was "One of these drawings is exactly square. Which drawing is it?" The answers simulated were 5,5,5,5.

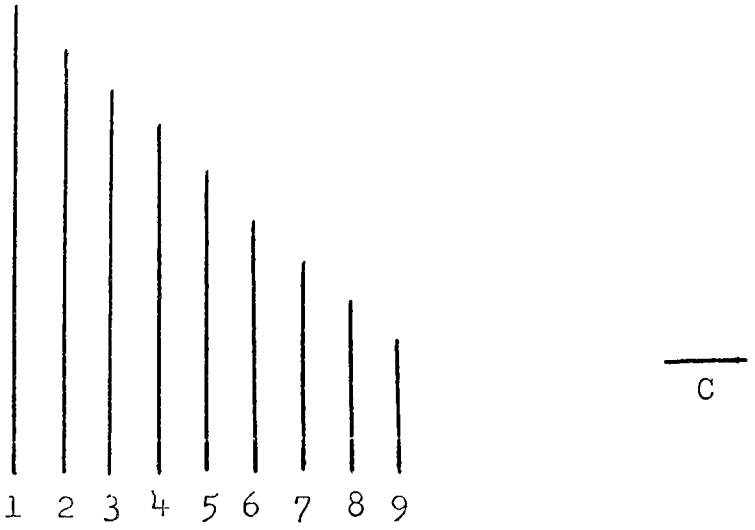


Figure 11.- Diagram of the Eighth Critical Item[a] in the Procedure for Measuring Conformity-Proneness.

Read D. Tuddenham, Studies in Conformity and Yielding X, Technical Report No. 11, Contract NR 170-159, Berkeley, University of California, 1960, p. 6.

a The accompanying verbal statement of the task was "One of the numbered lines can be cut into exactly three pieces the length of line C. Which line is it?" The answers simulated were 4,4,3,3.

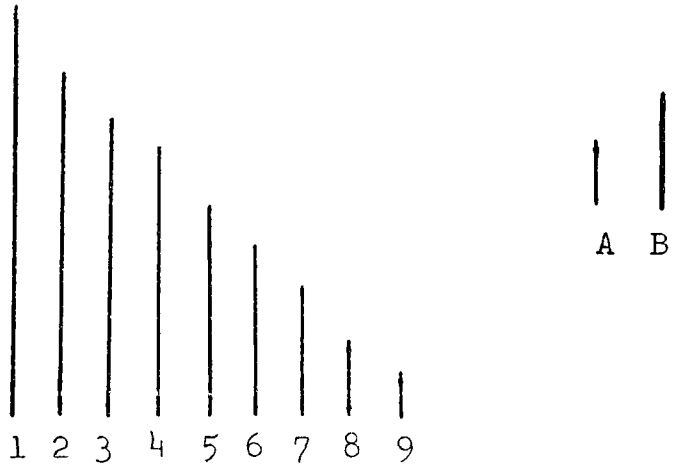


Figure 12.- Diagram of the Ninth Critical Item[a] in the Procedure for Measuring Conformity-Proneness.

Read D. Tuddenham, Studies in Conformity and Yielding X, Technical Report No. 11, Contract NR 170-159, Berkeley, University of California, 1960, p. 6.

a The accompanying verbal statement of the task was "One of the numbered lines is the same length as lines A and B joined end to end. Which line is it?" The answers simulated were 1,2,1,2.

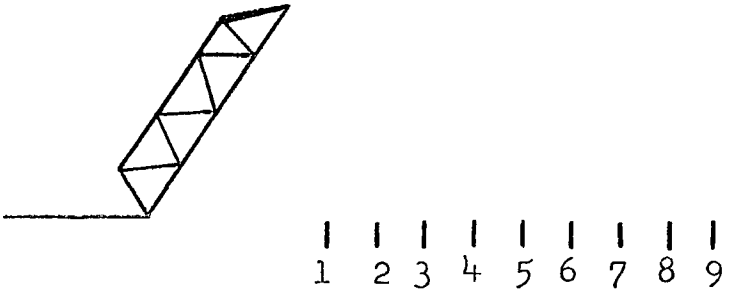


Figure 13.- Diagram of the Tenth Critical Item[a] in the Procedure for Measuring Conformity-Proneness.

Read D. Tuddenham, Studies in Conformity and Yielding X, Technical Report No. 11, Contract NR 170-159, Berkeley, University of California, 1960, p. 6.

a The accompanying verbal statement of the task was "If this bridge is lowered, what number will it reach?" The answers simulated were 7,7,7,7.

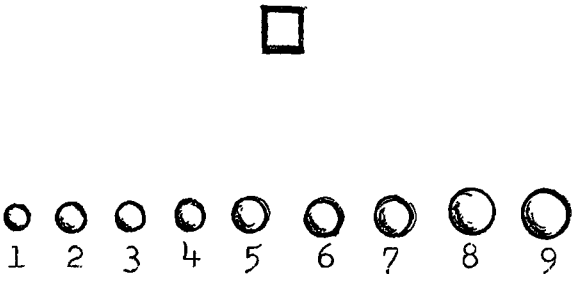


Figure 14.- Diagram of the Eleventh Critical Item[a] in the Procedure for Measuring Conformity-Proneness.

Read D. Tuddenham, Studies in Conformity and Yielding X, Technical Report No. 11, Contract NR 170-159, Berkeley, University of California, 1960, p. 6.

a The accompanying verbal statement of the task was "How many of the marbles are small enough to go through the square hole?" The answers simulated were 6,6,6,6.

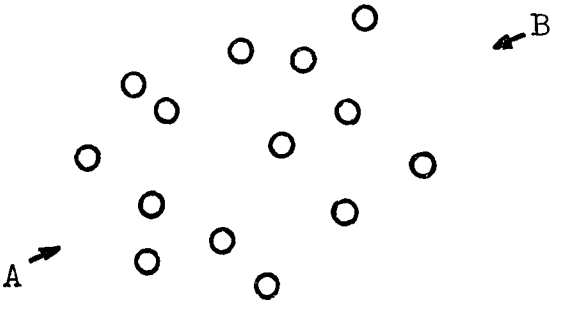


Figure 15.- Diagram of the Twelfth Critical Item[a] in the Procedure for Measuring Conformity-Proneness.

Read D. Tuddenham, Studies in Conformity and Yielding X, Technical Report No. 11, Contract NR 170-159, Berkeley, University of California, 1960, p. 6.

a The accompanying verbal statement of the task was "If you drew a line from Arrow A straight to Arrow B, how many little circles would it touch?" The answers simulated were 5,5,5,5.

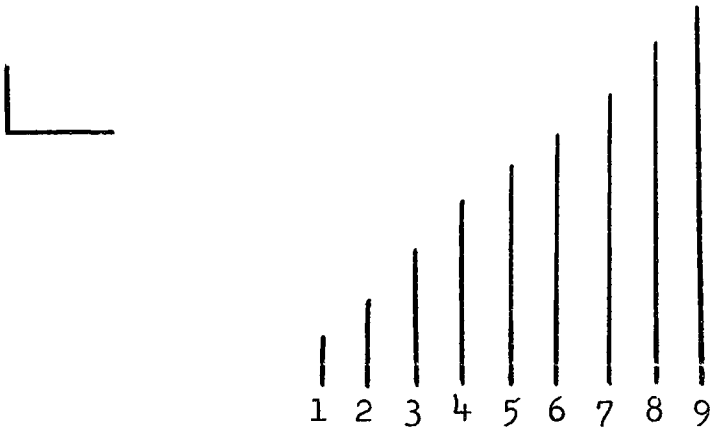


Figure 16.- Diagram of the Thirteenth Critical Item[a] in the Procedure for Measuring Conformity-Proneness.

Read D. Tuddenham, Studies in Conformity and Yielding X, Technical Report No. 11, Contract NR 170-159, Berkeley, University of California, 1960, p. 6.

a The accompanying verbal statement of the task was "Pretend the bent line is a piece of wire. When straightened out, it is the same length as one of the numbered lines. Which line?" The answers simulated were 8,9,9,8.

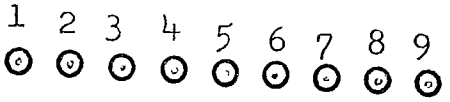


Figure 17.- Diagram of the Fourteenth Critical Item[a] in the Procedure for Measuring Conformity-Proneness.

Read D. Tuddenham, Studies in Conformity and Yielding X, Technical Report No. 11, Contract NR 170-159, Berkeley, University of California, 1960, p. 6.

a The accompanying verbal statement of the task was "This is a gun and some targets. Which target is the gun aimed at?" The answers simulated were 4,4,4,4.

APPENDIX 5

INSTRUCTIONS PLUS ANSWER SHEET FOR THE EC INVENTORY

TEST OF INSIGHT INTO OTHER PEOPLE

Name _____

Instructions

The statements in the booklet have been answered by a large group of females approximately your age. They described themselves by answering TRUE or FALSE to the statements. From their responses, the percentage of people answering TRUE was calculated for each statement.

The purpose of the present test is to measure how much insight you have into other people. You are required to judge what percentage of the group answered TRUE for each statement.

On this answer sheet there are two percentages listed for each statement. Circle the percentage which you think is the correct one; that is, the percentage of the large group of females who answered the statement TRUE about themselves.

Example A at the right has been marked that 25% of the people answered it TRUE.

	% TRUE	
	%	%
Example A	(25)	44

Give an answer for every statement. Make an intelligent guess when you are not sure.

	% TRUE			% TRUE			% TRUE			% TRUE			% TRUE	
	%	%		%	%		%	%		%	%		%	%
1.	9	23	25.	41	57	49.	1	7	73.	16	30	97.	24	38
2.	90	98	26.	32	46	50.	3	11	74.	47	63	98.	41	57
3.	37	53	27.	1	7	51.	89	97	75.	91	99	99.	59	73
4.	18	32	28.	6	14	52.	7	21	76.	1	7	100.	69	83
5.	26	40	29.	1	7	53.	1	7	77.	74	88	101.	20	34
6.	34	50	30.	89	97	54.	93	99	78.	66	80	102.	62	76
7.	50	64	31.	6	20	55.	78	92	79.	22	36	103.	90	98
8.	91	99	32.	23	37	56.	0	5	80.	25	39	104.	1	7
9.	95	100	33.	16	30	57.	65	79	81.	32	48	105.	82	96
10.	2	10	34.	6	14	58.	23	37	82.	18	32	106.	2	10
11.	1	9	35.	1	6	59.	36	52	83.	93	99	107.	93	99
12.	60	74	36.	66	80	60.	93	100	84.	15	29	108.	1	9
	% TRUE			% TRUE			% TRUE			% TRUE			% TRUE	
	%	%		%	%		%	%		%	%		%	%
13.	30	44	37.	89	97	61.	1	7	85.	0	7	109.	67	81
14.	2	10	38.	24	38	62.	21	35	86.	24	38	110.	1	9
15.	41	57	39.	31	45	63.	79	93	87.	33	49	111.	24	38
16.	0	7	40.	7	21	64.	27	41	88.	93	99	112.	55	69
17.	88	96	41.	32	46	65.	89	97	89.	20	34	113.	95	100
18.	78	92	42.	3	11	66.	4	18	90.	92	99	114.	1	6
19.	25	39	43.	1	9	67.	23	37	91.	28	42	115.	70	84
20.	93	99	44.	2	10	68.	80	94	92.	31	45	116.	23	37
21.	40	56	45.	67	81	69.	4	18	93.	26	40	117.	31	45
22.	12	26	46.	72	86	70.	62	76	94.	13	27	118.	6	14
23.	0	5	47.	1	9	71.	73	87	95.	36	52	119.	63	77
24.	1	7	48.	0	7	72.	1	7	96.	66	80	120.	75	89

APPENDIX 6

THE POST-EXPERIMENT QUESTIONNAIRE

Dear Volunteer Subject:

I have preliminary results from my research. It would be helpful for future studies if you would give your opinion on the theories I am investigating.

You Subjects answered two versions of a 120-statement inventory: one version was describing yourself, the other was a Test of Insight into the way other people answered the inventory. Often there was a similarity in the answering of the two versions. If a Subject marked a statement TRUE as applied to herself, often in the Test of Insight she judged that the larger percentage of other people would have answered it TRUE. And if a Subject answered FALSE to a statement as applied to herself, often in the Test of Insight she judged that the smaller percentage of other people would have answered it TRUE.

Perhaps Subjects were thinking in one of the following ways:

A--When in the Test of Insight a Subject was judging how other people answered the inventory, she might have expected other people to be like herself. If a statement was TRUE as applied to herself, she would expect many others to answer TRUE. So, she would circle the larger percentage in the Test of Insight. And in the inventory describing herself, she would answer the statement TRUE.

or

B--When answering the inventory describing herself, a Subject might have wanted to "go along with the majority." So, if she thought most people would answer a statement TRUE, she would mark it TRUE as applied to herself. Since she thought most people would answer the statement TRUE, she would circle the larger percentage in the Test of Insight.

Would you give your opinion about yourself. Which of the two ways of thinking, A or B above, was occurring when you were answering the inventories? If one was occurring more than the other in your case, please circle which one it was.

OPINION:

A was more often the way I was thinking.

B was more often the way I was thinking.

Please complete and return by tomorrow. Thank you.

Paul Sayer

APPENDIX 7

MEAN SR-EC CORRESPONDENCE SCORES FOR THE FACTOR COMBINATIONS

APPENDIX 7

Table VII.-

Mean SR-EC Correspondence Scores[a] for the Factor Combinations of the Analysis of Variance Design.

SD Set	Conformity	Order SR,EC	Order EC,SR
Low	Low	70.67	77.00
	Average	72.33	81.17
	High	71.33	75.33
Average	Low	77.67	71.33
	Average	72.83	72.67
	High	70.00	76.00
High	Low	71.33	77.33
	Average	73.17	76.17
	High	73.50	74.67

a Each score is the mean for six subjects.

APPENDIX 8

RAW DATA

APPENDIX 8

RAW DATA

Table VIII.-

Subjects' Scores for SD Response Set, Conformity, and SR-EC Correspondence.

Subject	SD Score[a]	Conformity Score[b]	Correspondence Score[c]
Low SD, Low Conformity, Order SR, EC			
S1	16	3	81
S2	19	10	62
S3	20	8	67
S4	23	5	66
S5	25	2	66
S6	26	3	82
Low SD, Low Conformity, Order EC, SR			
S7	19	4	83
S8	20	8	81
S9	20	9	90
S10	23	9	63
S11	25	3	68
S12	26	5	77
Low SD, Average Conformity, Order SR, EC			
S13	17	12	69
S14	21	12	80
S15	23	13	71
S16	24	17	58
S17	25	15	69
S18	27	18	87

Table VIII.- (Cont'd.)

Subject	SD Score	Conformity Score	Correspondence Score
Low SD, Average Conformity, Order EC, SR			
S19	17	16	80
S20	18	12	77
S21	19	17	86
S22	24	14	87
S23	25	12	78
S24	27	18	79
Low SD, High Conformity, Order SR, EC			
S25	14	28	71
S26	23	35	79
S27	24	18	74
S28	26	20	74
S29	27	21	67
S30	27	30	63
Low SD, High Conformity, Order EC, SR			
S31	15	20	77
S32	15	35	78
S33	23	25	76
S34	26	32	73
S35	27	19	65
S36	27	26	83
Average SD, Low Conformity, Order SR, EC			
S37	29	0	83
S38	29	2	84
S39	29	2	89
S40	29	6	86
S41	29	7	60
S42	31	7	64
Average SD, Low Conformity, Order EC, SR			
S43	28	7	75
S44	29	1	62
S45	29	6	76
S46	29	7	79
S47	30	2	80
S48	31	7	56

Table VIII.- (Cont'd.)

Subject	SD Score	Conformity Score	Correspondence Score
Average SD, Average Conformity, Order SR, EC			
S49	27	8	69
S50	28	10	74
S51	29	14	79
S52	30	15	72
S53	30	16	78
S54	31	13	65
Average SD, Average Conformity, Order EC, SR			
S55	28	8	76
S56	28	13	74
S57	29	10	75
S58	29	16	73
S59	30	10	64
S60	30	14	74
Average SD, High Conformity, Order SR, EC			
S61	27	23	86
S62	28	22	68
S63	29	31	69
S64	29	31	72
S65	30	28	66
S66	31	17	59
Average SD, High Conformity, Order EC, SR			
S67	28	19	70
S68	28	24	73
S69	28	28	75
S70	30	35	93
S71	31	21	75
S72	31	27	70

Table VIII.- (Cont'd.)

Subject	SD Score	Conformity Score	Correspondence Score
High SD, Low Conformity, Order SR, EC			
S73	32	7	78
S74	33	0	82
S75	33	1	79
S76	34	6	64
S77	34	8	66
S78	35	1	59
High SD, Low Conformity, Order EC, SR			
S79	32	0	71
S80	33	2	90
S81	33	4	91
S82	34	6	66
S83	35	1	72
S84	38	7	74
High SD, Average Conformity, Order SR, EC			
S85	32	13	73
S86	32	13	78
S87	33	10	80
S88	34	10	72
S89	34	15	79
S90	35	12	57
High SD, Average Conformity, Order EC, SR			
S91	32	12	74
S92	32	14	74
S93	32	16	77
S94	36	9	67
S95	36	11	76
S96	36	13	89
High SD, High Conformity, Order SR, EC			
S97	33	33	63
S98	34	27	77
S99	35	18	98
S100	35	20	55
S101	35	21	72
S102	37	42	76

Table VIII.- (Cont'd.)

Subject	SD Score	Conformity Score	Correspondence Score
High SD, High Conformity, Order EC, SR			
S103	32	34	66
S104	33	31	70
S105	34	30	90
S106	35	18	71
S107	35	21	78
S108	35	22	73

a The score was obtained from a subject's responses to Edwards' SD scale.

b The score represents the number of answer units by which a subject deviated from the norm in the direction of a bogus answer in the Crutchfield-Tuddenham procedure.

c The score represents the number of times a subject endorsed an SR item and answered the counterpart EC item with the overestimate, plus the number of times she answered an SR item FALSE and the counterpart EC item with the underestimate.

APPENDIX 9

ABSTRACT OF

The Influence of Social Desirability Response Set and of Conformity on the Correspondence Between Responses to Self-Report and Error-Choice Personality Inventories

APPENDIX 9

ABSTRACT OF

The Influence of Social Desirability Response Set and of Conformity on the Correspondence Between Responses to Self-Report and Error-Choice Personality Inventories[1]

The error-choice technique recently was proposed as a format of personality inventory which might eliminate distortion from social desirability response set. It requires subjects to judge the endorsement frequency of personality statements by choosing between two response options which are equidistant in opposite directions from the true endorsement frequency. The use of the technique as a method of personality assessment was predicated on the hypothesis that choosing the larger frequency in error-choice inventory would correspond to self-report endorsement.

A study provided some empirical evidence for the hypothesized correspondence, and the authors theorized that it might depend on a conformity motivation in self-report responses. The present investigation examined the influence of three factors on SR-EC correspondence: social desirability response set, conformity and order of administration of the inventories.

---

1 Paul E. Sayer, Master's thesis presented to the School of Graduate Studies of the University of Ottawa, Ontario, 1972, xi-126 p.

The subjects were 108 female high school graduates. They were classified according to social desirability response set by means of Edwards' SD scale, and according to conformity-proneness by means of Crutchfield's laboratory procedure. A self-report inventory and an error-choice version were administered, half the subjects receiving them in one order and half in the reverse order.

The first hypothesis stated that there are no significant differences in the amounts of SR-EC correspondence shown by subjects who are classified as low, average and high in SD response set. No significant differences were obtained, and the hypothesis was not rejected. Results were interpreted as supporting the theory that error-choice inventory eliminates the influence of SD response set.

The second hypothesis stated that there are no significant differences in the amounts of SR-EC correspondence shown by subjects who are classified as low, average and high in conformity. No significant differences were obtained, and the hypothesis was not rejected. The lack of significant findings did not support the rationale that a conformity motivation in self-report inventory accounts for SR-EC correspondence.

The third hypothesis stated that there is no significant difference in the amounts of SR-EC correspondence shown by subjects who answer the SR inventory followed by the EC

version, and by subjects who answer the EC inventory followed by the SR version. The data did not yield significant differences, and the hypothesis was not rejected. However, a test of the simple effects of orders of administration for subjects low in SD set approached significance. Consequently, it was suggested that the order variable should be controlled in subsequent research.

The fourth hypothesis stated that there are no significant interactions among the three factors. No significant differences were observed, and the hypothesis was not rejected.

With regard to the failure to find significant results, the low powers of the tests of significance were noted. In addition, the construct validity of the measuring instruments was discussed.

In a post-experiment questionnaire, the large majority of subjects reported that they sometimes estimated endorsement frequencies by conceiving others in their own image. That was interpreted as tentative support for the theory that error-choice inventory reveals the personality of the respondent.

The hypothesized SR-EC correspondence occurred beyond chance expectations, and the error-choice format of personality inventory was judged to be worthy of further research.